

# THIS WEEK



## EDITORIALS

**WORLD VIEW** Jordan drafts stem-cell rules under Islamic law **p.189**

**AGRICULTURE** Spider venom pesticide found safe for bees **p.190**

**COGNITION** Visual memory could give crows their bird brains **p.191**

## A growing problem

*Without careful stewardship, genetically engineered crops will do little to stop the spread of herbicide-resistant weeds.*

**P**almer pigweed (*Amaranthus palmeri*) is not a weed to trifle with. It can reach more than 2.5 metres tall, grow more than 6 centimetres a day, produce 600,000 seeds and has a tough, woody stem that can wreck farm equipment that tries to uproot it.

It is also becoming more and more resistant to the popular herbicide glyphosate.

The first such resistant population was confirmed in 2005 in a cotton field in Georgia, and the plant now plagues farmers in at least 23 US states. It is just one of many resistant weeds marching through the world.

The US Environmental Protection Agency (EPA) is trying to learn from the pigweed experience, and wants to limit the damage caused by the latest wave of weed control. It deserves credit and support.

There is broad agreement that the spread of these resistant plants has its roots in the widespread adoption of crops engineered to be resistant to glyphosate. By the time these genetically engineered crops were released in the mid-1990s, farmers had been battling herbicide-resistant weeds for decades. But glyphosate was thought to be a particularly challenging herbicide for weeds to overcome. Few cases of resistance had been seen.

That was set to change: by 2012, glyphosate-resistant weeds had infested 25 million hectares of US cropland. They have also appeared in other countries that have embraced glyphosate-tolerant crops, including Australia, Brazil and Argentina. Blanketing crops year after year in the same herbicide is the perfect way to foster resistant weeds.

Chemical companies have come up with a solution: crops engineered to tolerate multiple herbicides. The likelihood of a weed becoming resistant to more than one chemical, they claim, is very small. And, in an eerie echo of the 1990s discussion around glyphosate tolerance, some even point out that one of the other herbicides being targeted — the choline salt of an old chemical called 2,4-D — has been used for decades with little sign of resistance.

It is a flawed argument. Stacking up tolerance traits may delay the appearance of resistant weeds, but probably not for long. Weeds are wily: farmers have already reported some plants that are resistant to more than five herbicides. And with glyphosate-resistant weeds already in many fields, the chances of preventing resistance to another are dropping.

Crops resistant to multiple herbicides could be useful. But scientists are concerned that farmers will rely too heavily on the chemicals, and neglect other ways to combat the resistance threat. Those include using a mixture of herbicides that are specific to a field's invaders, rotating crops and moderate tilling — practices together known as integrated weed management. A farmer making good money in the age of biofuel crop subsidies may be loath to switch to a different crop. And farmers may be hesitant to invest the money needed to properly manage weeds, when their farms could end up infested with weeds from less-assiduous neighbours.

This is where the EPA comes in. In its draft assessment of the

blend of herbicides to be used, it calls for the manufacturer — Dow AgroSciences of Indianapolis, Indiana — to monitor the emergence of resistant weeds and report them to the agency. The EPA will then have the power to impose restrictions on Dow or on the use of the herbicide if it deems this necessary.

The EPA is soliciting comments on the draft assessment from the public until the end of June. It offers sensible precautions, but it could

**“The EPA proposes sensible precautions, but it could do much more.”**

do much more. When an insect-resistant variety of genetically engineered crop was released, US regulators required farmers to plant nearby refuges of non-resistant plants to ease the selection pressure on insects to develop resistance to the crops. Similar measures for herbicide-tolerant crops might require farmers to rotate crops or herbicides every few years — a familiar restriction, because many herbicides have limits on how often they can be

used for environmental reasons. Such measures would be a sign that regulators and farmers alike have realized the consequences of underestimating the ability of weeds to develop resistance. ■

## Good practice

*Standardized procedures and analyses should help to get stem-cell therapies to the clinic.*

**U**nethical procedures, exploitation and inflated promises, that is what generally makes the headlines — and so it is with regenerative medicine and stem cells. Media reports have left the distinct impression that the research is rather dubious.

First is the long-standing controversy over the source material: human embryos. Research banned by the most powerful man in the world — as US President George W. Bush was when he stopped federal support for such work in 2001 — must be a bit dodgy, right? Then there are the regular reports of companies that are exploiting vulnerable — and often seriously ill — patients with promises of expensive, but unproven, miracle cures.

But behind the headlines is a different story. Scientists doing the systematic research needed to get cellular therapies into the clinic are finally making headway. Trials are now under way for treating an eye disorder called macular degeneration using retinal cells. And a trial using immature glial cells to treat spinal-cord injury has restarted after the company running it pulled out in 2011 (see *Nature* **510**, 18; 2014).

It has taken many years to get to the starting line, but shortcuts are simply not possible, despite charlatan claims. It takes time to

learn how to coax stem cells — either from human embryos or from reprogrammed adult cells known as induced pluripotent stem (iPS) cells — to develop into the right sort of replacement cell. It also takes time to work out how to get these cells to integrate into the host tissue and to function. And the steps required to work out how many replacement cells need to be delivered, and how to deliver them safely, cannot be rushed.

The eye and spinal cord are relatively isolated systems. Much will be learnt from them, but the brain and heart are altogether more complicated. Fixing damage in these systems is crucial, however, because together they provide the biggest disease burden in developed countries.

Happily, clinical trials are on the horizon. Treatments for Parkinson's disease are just a few years away from clinical testing. And some for Huntington's disease may not be far behind.

Taking any radical therapy into humans requires caution. Ideally, researchers should be able to use data from a patient in one trial to refine the approach for one in another. So a decision by the Global Force for Parkinson's Disease, or G-force, to bring together teams from Europe, the United States and Japan to define standards for cell preparation and patient selection and monitoring for future trials is particularly welcome (see page 195).

The G-force seems to have learnt the lessons of moving research to the clinic too fast and in isolated teams. Multiple trials of cells derived from fetal brains to treat Parkinson's disease began in the late 1980s, but stopped in 2003 because the outcomes were an uninterpretable mishmash. And trials using adult stem cells to treat heart failure have shown wildly varying outcomes (see *Nature* **509**, 15–16; 2014),

perhaps owing in part to a lack of good preclinical data. But systematic research has now shown that heart cells derived from human embryonic stem cells can engraft into damaged primate hearts and synchronize their beats to it, at least to some extent. Some of the monkeys developed arrhythmias, showing that the technique still needs improvement. The principle of the therapy has been proven, however, which gives confidence that clinical trials may become possible.

**“News reports need to be careful not to overhype the potential of cellular therapies.”**

Designing trials to agreed standards will ensure that researchers can understand why any one patient benefited, or failed to benefit, from the treatment. This will magnify the efficiency of the trials and speed up the development of therapies. It is a model that deserves to be widely copied.

News reports need to be careful not to overhype the potential of cellular therapies. As the field inches towards clinical testing, it is important that researchers clearly communicate to the media what the therapies are likely to achieve — and what they are not. Early trials are unlikely to show cures, but that does not diminish their value: even small improvements in quality of life are important to a person with a serious disability. A blind person who becomes able to discern light from shade, a paralysed person who regains some feeling in a limb and a person with advanced Parkinson's disease who can walk independently, if not normally — each will think it worthwhile.

Like all new therapies, stem-cell repair will improve through trial and error. These approaches promise more trial and, hopefully, fewer errors. ■

## Open goal

*International researchers can help to improve the scientific enterprise in South America.*

**P**roductivity in offices and labs around the world will probably slip a little during the next month, as football fans tune in to watch the 2014 FIFA World Cup, which starts in Brazil this week. Four years ago, nearly half the world's population tuned in at some point during the tournament. And as the world focuses its attention on Brazil, *Nature* has taken the opportunity to widen the view with our special issue on science in South America (see page 201). The package of articles and commentaries details some of the success stories on the continent as well as the substantial challenges faced by researchers there as they seek to build scientific institutions in the wake of decades lost to dictatorships.

They need not struggle alone. From London to Boston to Tokyo, individual scientists and larger organizations in the developed world can offer significant help to South American countries. When *Nature* asked leading South American scientists what kind of assistance would bring tangible benefits, the answers invariably clustered around two key requests to their international colleagues: host young scientists in your laboratories, and come to visit South American researchers.

The flow of students from South America to the United States and Europe has grown in recent years but remains a trickle. Brazil sent fewer than 11,000 undergraduate and graduate students to the United States in 2013 — less than Turkey and Vietnam, countries with much smaller populations and economies. The tally for all students sent to US universities from Latin America and the Caribbean was less than one-third of the number sent by China.

Many South American scientists called on their northern colleagues to recruit more graduate students and postdoctoral

scientists from the continent. Even short visits of three to six months can help to train a young scientist. But the exchanges have to be done in a way that does not contribute to the brain drain that has lured many leading researchers to permanent positions in the United States and Europe (see pages 207 and 213). One solution is to provide start-up funds for researchers returning to South America. For example, after postdoctoral training in the United States, Lino Barañao received support from the Rockefeller Foundation to establish his lab at home in Argentina, where he is now the minister of science, technology and innovative production.

Travel needs to go both ways. According to South American researchers, too few scientists visit their continent to spend time in labs, give lectures and attend meetings. Even virtual visits, through video conferences, would help.

The networking requests go beyond the wish to trade research methods and results. Scientists in South America want to know how to select the best people and how to improve coordination between universities and industry. Many called for help in improving science-evaluation processes (see page 209). In Brazil, for example, assessments too often reward quantity over quality.

Investments in sending researchers back and forth can yield long-term dividends. In 1990, Argentine molecular biologist Eduardo Arzt started a fellowship at the Max Planck Institute for Psychiatry in Munich, Germany. After returning to Argentina, Arzt continued to collaborate with Max Planck colleagues — a connection that was key when the society was looking to expand its international programs. In 2011, it established its first South American partner institute in Buenos Aires, run jointly with Argentina's Council for Scientific and Technological Research, and with Arzt as director. Several of the research groups at the institute are led by Argentine scientists lured back from overseas by the opportunity to do top-tier science.

**➔ NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunjv](http://go.nature.com/xhunjv)

Football fans in South America are used to seeing top players leave for abroad. Efforts to reverse the flow, in science as in sport, face great challenges. But they are a worthwhile goal. ■



A. AWAD



## Jordan's stem-cell law can guide the Middle East

A ban on private companies using stem cells from human embryos provides a policy framework for other Arab and Islamic countries, says **Rana Dajani**.

In January, Jordan passed a law to control research and therapy using human stem cells derived from embryos — the first such regulation in the Arab and Islamic region. I was part of the group headed by Abdalla Awidi Abbadi, director of the Cell Therapy Center at the University of Jordan in Amman, that initiated the call for the law and later drafted it. Stem-cell research is a hot topic for Jordan because of the kingdom's status as a health-care hub that draws patients from abroad. It is already one of few countries in the Middle East with regulations for protecting people who participate in clinical trials. This latest law should serve as an example to other countries in the region.

The new rules ban private companies from using human embryonic stem (ES) cells in research or therapies. Such work will be allowed only in government organizations or publicly funded academic institutions in Jordan, which have higher levels of transparency than private firms and are supervised by the health ministry and a specialized committee. The law also bans payment for donations of stem cells and eggs, and says that modified and manipulated cells are not to be used for human reproduction. There is no current research on human ES cells in Jordan; this is a pre-emptive step.

Much of the controversy and disagreement over work on stem cells worldwide arises from the different views of the major religions on the earliest stages of life. Although the use of human ES cells is opposed by the Roman Catholic Church and some Protestant denominations, it is generally supported by the Jewish community and accepted in many Muslim countries. There is no consensus on when human embryonic life begins, but the majority of Muslim scholars consider it to start 40–120 days after conception and therefore hold the view that a fertilized egg up to 5 days old has no soul — it is not 'human life' but 'biological life'. So for many, there is no ethical problem in the Islamic faith with using an early embryo to produce stem cells.

Such conclusions are not easy to reach. Many Muslim countries consider legislation and bioethics principles to be based on three pillars of Islamic law. The first is the Quran. The second is Sunnah, or the legislative decisions of the Prophet Muhammad. The third is *ijmaa* — the consensus of Muslim scholars — and *ijtihad*, the concept that every adequately qualified scholar has the right to independently solve problems. On the basis of these pillars, Iran, Saudi Arabia and Tunisia have drawn up guidelines on stem-cell research, but they are not legally binding.

Jordan's stem-cell law is the product of years of discussions by committees comprising scientists, physicians, Arabic-language experts, lawyers and Muslim and Christian theologians. The issues that arose — confusion between 'stem cells' and 'embryonic stem cells', for instance — were discussed and

resolved. We consulted with both the National Committee for Science and Technology Ethics and the education ministry. The final law was approved by the council of Muslim scholars, the Majlis Al-Iftaa.

The council agreed with a 2003 decision (*fatwa*) by Muslim scholars that allows the use of human ES cells from permissible sources — including legally produced excess fertilized eggs from *in vitro* fertilization. The decision to ban private companies from using these cells was driven by concerns that the work would encourage termination of pregnancies, which is illegal in Jordan unless the mother's life or health is at risk. The council was clear that the new law must forbid human reproductive cloning and should not allow embryos to be created from the sperm and eggs of unmarried couples.

The distinction drawn between the various sources of stem cells earlier in the discussion process allowed the Majlis Al-Iftaa to take a more permissive approach to techniques using stem cells that are not derived from human embryos. For example, somatic-cell nuclear transfer (in which a patient's DNA is transplanted into an unfertilized human egg that has no nucleus) and induced pluripotent stem cells, which are made from adult cells, can be worked on by the private sector under the new rules.

The therapeutic use of bone-marrow transplantation — including transplants of blood-forming stem cells — is well established in Jordan. Such procedures are already regulated by existing laws on medical practice, so the new law makes a clear distinction between these techniques and human ES-cell therapy.

The legislation not only covers all current aspects of stem-cell research and use, but also leaves room for later modification. It mandates

the creation of a national committee that, among other things, will take responsibility for laying out specific regulations for stem-cell banking in accordance with international standards.

All our discussions in Jordan have concluded that stem-cell research is permissible in Islam, as long as it is carried out to improve human health and takes precautions to respect human life. Still, as the field develops, policy-makers must continue to invest in education and raise awareness of the opportunities, challenges and uncertainties of human ES-cell research.

The scientific output of the Islamic Arab region is low compared with that of other regions. Implementation of these laws in Jordan and other Muslim countries could help to encourage research to reach international standards and start to bridge that divide. ■

**Rana Dajani** is associate professor of molecular cell biology at the Hashemite University in Zarqa, Jordan.  
e-mail: [rdajani@hu.edu.jo](mailto:rdajani@hu.edu.jo)

ALL OUR  
**DISCUSSIONS  
IN JORDAN  
HAVE CONCLUDED  
THAT STEM-CELL  
RESEARCH IS  
PERMISSIBLE  
IN ISLAM.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/anosjp](http://go.nature.com/anosjp)

## CANCER

## Immune cells targeted in cancer

In a small, early-stage clinical trial, an antibody seems to slow the growth of tumours by decreasing the number of cancer-boosting immune cells in and near the tumours.

Some immune cells known as macrophages promote tumour growth and are regulated by a protein, CSF-1, and its receptor. Carola Ries at Roche in Penzberg, Germany, and her colleagues produced an antibody that blocks this receptor and tested it in seven patients with a rare cancer of the joints. The researchers found that the antibody lowered the number of macrophages in one patient from whom a biopsy was taken, and shrank tumours in five of the patients. In people with other types of tumours, the antibody also depleted tumour-associated macrophages and shifted the ratio of another type of immune cell, T cells, towards those that fight tumours.

Targeting macrophages, in combination with other chemo- or immunotherapies, could improve treatment, but further testing in humans is needed, the authors say.

*Cancer Cell* <http://doi.org/s3m> (2014)

## BIOCHEMISTRY

## Easy monitoring of drug by camera

By adding a drug-sensing molecule to human blood samples, researchers can measure drug levels with a simple digital camera.

Monitoring drug amounts in patients can avoid side effects, but the process requires specialized resources. Now, Kai Johnsson at the Swiss Federal Institute of Technology in Lausanne and his colleagues

have used a digital camera and software to quantify blood levels of a cancer drug bound to a specially designed bioluminescent sensor protein. The sensor, which changes from red to blue (**pictured**) with increasing drug levels, can be tailored to other drugs, and could allow easy, low-cost drug monitoring by physicians and patients, the authors say.

*Nature Chem. Biol.* <http://doi.org/s5b> (2014)

## IMMUNOLOGY

## Skin sensor soothes psoriasis

A protein in the skin that senses environmental signals could be enlisted to fight inflammation caused by the autoimmune skin disease psoriasis.

Brigitta Stockinger at the MRC National Institute for Medical Research, London, and her team found that altering the activity of the protein AhR in human skin affected the expression of 41 genes that are relevant to psoriasis.

Mice lacking AhR had a stronger response to imiquimod, a compound that causes psoriasis-like skin inflammation. However, stimulating AhR in normal mice reduced imiquimod's effects, suggesting that AhR activation may ease psoriasis. *Immunity* <http://doi.org/s4m> (2014)

## PHYSICS

## Another source for static electricity

Physicists have debunked a three-decades-old explanation for how grains of the same material rub together to generate static electricity — an effect seen, for example, in volcanic ash clouds.

One theory posited that because larger grains hold more trapped, high-energy electrons,

## SOCIAL SELECTION

Popular articles on social media

## Papers predict future lab heads

Scientists at every point on the career spectrum are talking about a paper in *Current Biology* that takes a quantitative view of the mantra 'publish or perish'.

Using a sample of more than 25,000 researchers, Lucas Carey at Pompeu Fabra University in Barcelona, Spain, and his colleagues developed a statistical model that predicts who will eventually become principal investigators. The team found that first authors of papers in high-impact journals have the inside track, and everyone else is likely to lag behind. Verena Seufert, a geographer and PhD candidate at McGill University in Montreal, Canada, tweeted that it was a "sad story from a cool paper".

Van Dijk, D., Manor, O. & Carey, L. *Curr. Biol.* 24, R516–R517 (2014)



Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

➔ **NATURE.COM**  
For more on popular papers:  
[go.nature.com/meav9t](http://go.nature.com/meav9t)

they redistribute the electrons to smaller grains when two touch, creating static electricity. Heinrich Jaeger from the University of Chicago, Illinois, and his colleagues measured the surface density of trapped electrons around different-sized grains of zirconium dioxide silicate, as well as the grains' charge. The authors found that there are far too few trapped electrons to account for the observed static build-up when the grains are mixed.

Instead, other charged particles, such as ions from water films or from the surrounding atmosphere, could accumulate on the grains' surface and be responsible for the effect, the team suggests.

*Phys. Rev. Lett.* 112, 218001 (2014)

## ANIMAL COGNITION

## Crow brain recalls images

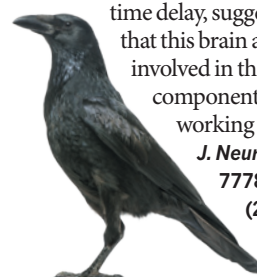
The exceptional cognitive abilities of crows could be partly due to a structure in their

brain that can temporarily retain visual information.

To see whether crows have aspects of working memory — the ability to remember information for future tasks — Andreas Nieder and his team at the University of Tübingen, Germany, trained four carrion crows (*Corvus corone*, **pictured**) in a task that required them to recall images 1 second after first seeing them. During this task, the team recorded the activity of 662 individual neurons in a region of the brain called the nidopallium caudolaterale, which is thought to correspond to the mammalian prefrontal cortex — an area involved in higher-order thought.

The neurons seem to encode and maintain information about the image during this time delay, suggesting that this brain area is involved in the visual component of working memory.

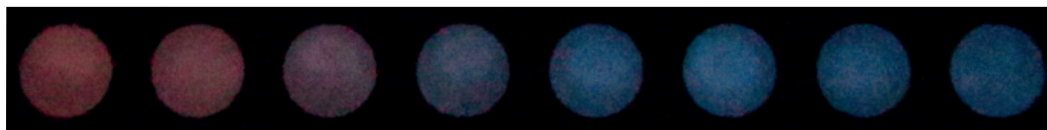
*J. Neurosci.* 34, 7778–7786 (2014)



➔ **NATURE.COM**  
For the latest research published by Nature visit:  
[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

LIFE ON WHITE/PHOTODISC/GETTY

RUDOLF GRISSE/EPFL



# SEVEN DAYS

The news in brief

## RESEARCH

### Destination Mars

To revive the moribund US human spaceflight programme, NASA should plot a course for Mars, says the US National Academy of Sciences in a report published on 4 June. The report outlines plans for putting humans on Mars sometime between 2037 and 2050 at a cost of hundreds of billions of dollars. It criticizes the agency's current strategy of seeking to visit an asteroid put in orbit around the Moon, warning that it will end in the "loss of the long-standing international perception that human spaceflight is something the United States does best". See [go.nature.com/wp4zdl](http://go.nature.com/wp4zdl) for more.

### Forest genetics

Many countries do not know enough about the genetic make-up of the trees growing in their native forests, finds a report from the Food and Agricultural Organization of the United Nations (FAO). The 86 countries that contributed to the analysis provided genetic information for only around 600 species out of a maximum of 100,000 shrubs and trees thought to be growing around the globe, the report says. In its analysis, *The State of the World's Forest Genetic Resources*, published on 3 June, the FAO calls on governments to improve data gathering and research to help manage tree species.

### El Niño watch

The US National Oceanic and Atmospheric Administration predicts that there is an 80% chance that an El Niño event — a periodic warming of waters in the eastern equatorial Pacific Ocean — will occur this autumn or winter. The forecast, made on 5 June, also predicts a 70%



## Nature in close-up

This microscopic image shows the forest-like arrangement of hairs on a gecko's toe that gives the animal its gravity-defying ability to scurry across ceilings. Each foot has hundreds of thousands of these hairs, called setae, which fray into smaller hairs with split ends called spatulae. The hairs' strong grip has inspired the design of medical adhesives. This image was taken by Dennis Kunkel, a photomicrographer in Hawaii. It is part of *Life: Magnified*, an exhibition of scientific images on display at Washington Dulles International Airport's Gateway Gallery from June to November. Other images include a bacterium being swallowed by an immune-system cell and chromosomes lining up for cell division.

chance that an El Niño will occur this summer — up from a 50% chance predicted in March. The US forecasters suggest a moderate-strength El Niño, which could scramble global weather patterns until it ebbs.

## FACILITIES

### SOFIA saved

The world's biggest flying telescope, the Stratospheric Observatory for Infrared

Astronomy (SOFIA), was given a new lease of life on 5 June. The US Senate voted to give US\$87 million in the 2015 fiscal year to the observatory — a modified Boeing 747 that carries a 2.5-metre telescope. The funding boost could rescue SOFIA; in March, NASA had proposed effectively cancelling the project because of its high operating costs. SOFIA is a joint venture with the German Aerospace Center and became

fully operational in February. The Senate and the House of Representatives must now agree on the budget.

### Telescope pull-out

Germany is pulling out of the world's largest radio telescope, the Square Kilometre Array (SKA), scheduled to be completed in South Africa and Australia by the mid-2020s. Germany's research ministry announced the move on 5 June, citing a tight budget, according to the SKA Organisation. The decision will take effect on 30 June 2015. The pull-out is "disappointing, but not catastrophic" for SKA's ability to secure funding, says Philip Diamond, director-general of the SKA Organisation near Manchester, UK. See [go.nature.com/cuacno](http://go.nature.com/cuacno) for more.

## BUSINESS

### Stem-cell patents

A US federal court threw out a legal challenge to a key embryonic-stem-cell patent on 4 June. The non-profit advocacy group Consumer Watchdog in Santa Monica, California, had argued that the patent was invalid because the supposed invention was merely a product of nature. The US Court of Appeals for the Federal Circuit ruled that Consumer Watchdog could not challenge the patent, which is owned by the Wisconsin Alumni Research Foundation in Madison, because the group does not use embryonic stem cells and was not directly harmed by the patent.

### Pharma buy-out

US pharmaceutical giant Merck agreed to pay US\$3.85 billion for Idenix — a developer of hepatitis C virus therapies in Cambridge, Massachusetts. The price tag,

DENNIS KUNKEL/DENNIS KUNKEL MICROSCOPY/NIH/MS



ANTHONY PIDGEON/REDFERNS/GETTY

agreed on 9 June, was more than three times the Idenix stock listing at the close of trading on the previous day.

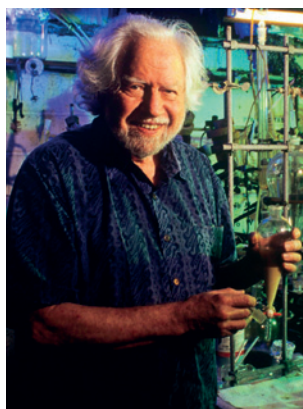
## PEOPLE

## Fossil sentence

An American fossil dealer was sentenced to three months imprisonment on 3 June, after pleading guilty to smuggling dinosaur fossils into the United States, including a 70-million-year-old *Tyrannosaurus bataar* skeleton from Mongolia. Eric Prokopi received a reduced sentence for helping prosecutors to recover more than 17 other dinosaur fossils. The stolen *Tyrannosaurus* skeleton, which sold at auction for more than US\$1 million, was returned to Mongolia in May 2013. Mongolia will open a dinosaur museum to house the returned fossils.

## Chemist dies

Alexander 'Sasha' Shulgin, a chemist famed for synthesizing and analysing novel psychoactive compounds, died on 2 June at the age of 89. Shulgin (pictured) earned his PhD in biochemistry from the University of California, Berkeley, and went on to work for US firm Dow Chemical, where he developed one of the first successful biodegradable pesticides, mexacarbate.



Later he began to develop psychedelic drugs and tested their activity on himself and a small group of friends. He published numerous academic studies and books on the subject, and remained a respected scientist throughout his career.

## POLICY

## BRAIN plan

The US National Institutes of Health laid out its ten-year plan for the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative on 5 June. It proposed that Congress grant the initiative a further US\$4.5 billion for 2016–25. According to the plan, the first five years will be spent developing technologies to record, analyse and manipulate the brain. In the following five years,

researchers will use those technologies to study how the brain's circuits lead to behaviour and cognition.

## High-seas value

International waters store around 500 million tonnes of atmospheric carbon per year, providing an 'ecosystem service' to humans that is worth up to US\$222 billion annually, according to the first economic assessment of the high seas. The report was published on 5 June by the Global Ocean Commission, a non-governmental organization in Oxford, UK. It adds that 10 million tonnes of fish are caught each year in international waters, generating more than \$16 billion. On 24 June, the commission will publish proposals for protecting the ocean.

## Microbead ban

Illinois became the first US state to ban the manufacture and sale of personal-care products that contain plastic microbeads. Environmental scientists say that the non-biodegradable beads, used as exfoliating agents, pass through sewage systems and build up in waterways, where they absorb toxic chemicals and threaten aquatic life. A law signed by Governor Pat Quinn on 8 June prohibits the manufacture of soaps, cosmetics and

## COMING UP

## 17–19 JUNE

Space scientists meet in Chicago, Illinois, for the annual research conference on the International Space Station.

[go.nature.com/wnatte](http://go.nature.com/wnatte)

## 19 JUNE

The peak of the 3,000-metre Cerro Armazones mountain in northern Chile will be blown off to make a home for the European Extremely Large Telescope.

[go.nature.com/u6xrrb](http://go.nature.com/u6xrrb)

medications containing the beads by 2018 and the sale of these products by 2019.

At least four other states are considering similar bills.

## Carbon cap

China could set its first absolute cap on fossil-fuel emissions from 2016. On 3 June, international media reported that He Jiankun, a senior government adviser on climate change, told a meeting in Beijing that the Chinese government may outline the cap in its next five-year economic plan, for 2016–20. He later clarified that the idea was his personal view. China is the world's biggest emitter of carbon dioxide.

## Development goals

A United Nations working group led by Hungary and Kenya has drawn up a list of 17 sustainable development goals and started negotiating on them this week. The goals will replace the Millennium Development Goals, which expire next year and include targets for eradicating hunger and poverty and halting biodiversity loss. The United Nations aims to finalize the new goals next year.

► [NATURE.COM](http://NATURE.COM)

For daily news updates see:

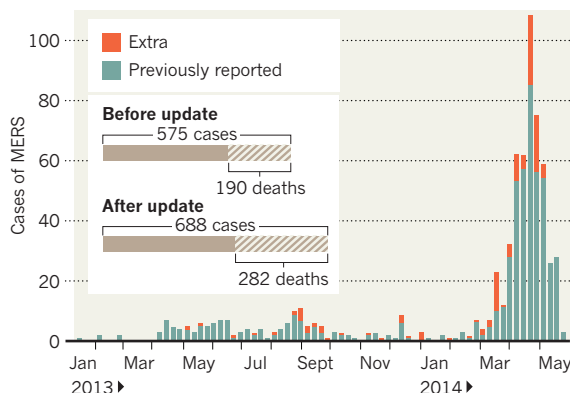
[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

The fatality rate for cases of the Middle East respiratory syndrome (MERS) coronavirus in Saudi Arabia is 41%, not 33%, according to figures released on 3 June. The country's health ministry — which a day earlier sacked deputy health minister Ziad Memish, a key figure in the nation's efforts to contain the virus — said that it had retrospectively identified 113 extra cases, and announced "new standards" for reporting the disease. Of 815 MERS cases reported worldwide by 4 June, 84% were in Saudi Arabia.

## SAUDI ARABIA FINDS NEW CORONAVIRUS CASES

Health ministry reports retrospective discovery of 113 extra cases of Middle East respiratory syndrome (MERS) — of whom 92 died.



SOURCE: MINISTRY OF HEALTH, SAUDI ARABIA

# NEWS IN FOCUS

**SPACE** Uncertainty over future triggers push in space-station research **p.196**

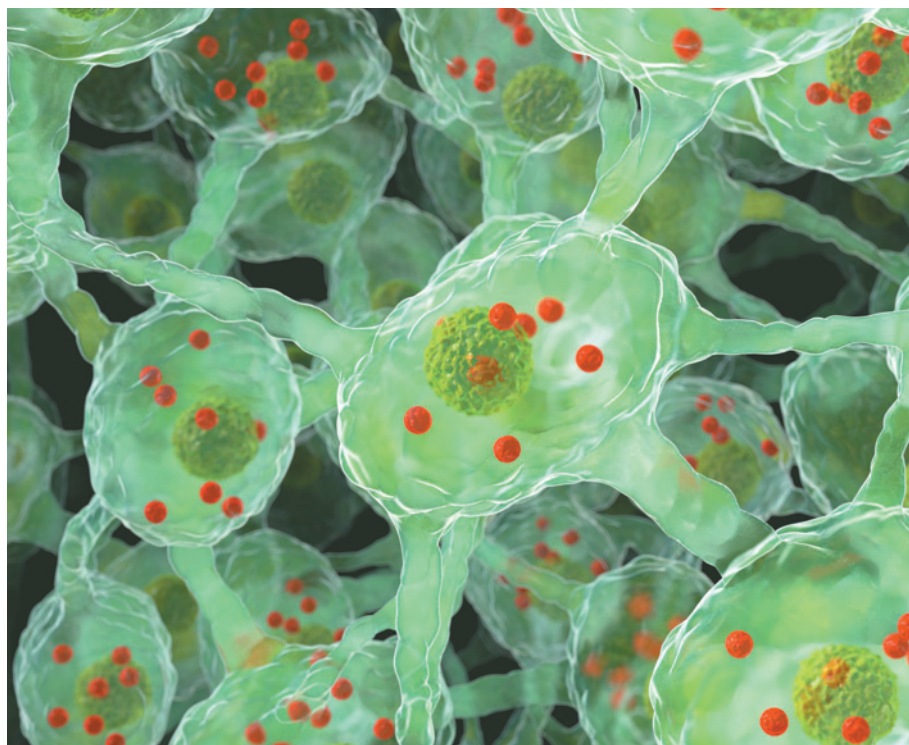
**GENETICS** Competition over breast-cancer tests heats up a year after patent ban **p.198**

**ACOUSTICS** Marine signal offers lead in hunt for missing plane **p.199**



**SOUTH AMERICA** Seeds of success build hope for stronger science **p.201**

ANIMATED HEALTHCARE LTD/SPL



Defective brain neurons are responsible for the mobility problems seen in people with Parkinson's disease.

## REGENERATIVE MEDICINE

# Fetal-cell revival for Parkinson's

*Moratorium on controversial therapy lifted as stem cells emerge as alternative source of treatment.*

BY ALISON ABBOTT

A neurosurgery team will next month transplant cells from aborted human fetuses into the brain of a person with Parkinson's disease. The operation breaks a decade-long international moratorium on the controversial therapy that was imposed after many patients failed to benefit and no one could work out why.

But the trial comes just as other sources of

replacement cells derived from human stem cells are rapidly approaching the clinic. And this time, scientists want to make sure that things go better. So the teams involved in all the planned trials have formed a working group to standardize their research and clinical protocols in the hope that their results will be more easily interpretable.

People with Parkinson's disease suffer from a degeneration of neurons that produce the neurotransmitter dopamine, which is crucial for

normal movement. This often leaves patients with severe mobility problems. Standard treatment includes the drug L-dopa, which replaces dopamine in the brain but can cause side effects. The cellular therapies aim to replace the missing neurons with dopamine-producing (dopaminergic) cells from fetal brains or with those derived from human stem cells.

The moratorium on replacement-therapy trials was introduced in 2003 because the early fetal-cell studies had produced varying results that were impossible to interpret.

"We want to avoid a repeat of this situation," says neurologist Roger Barker at the University of Cambridge, UK, who helped to organize the working group's inaugural meeting in London last month. The group, known as the Parkinson's Disease Global Force, includes scientists from the European, US and Japanese teams about to embark on the trials. At the meeting, they pledged to share their knowledge and experiences.

The first human transplantation of fetal brain cells took place in 1987 at Lund University in Sweden, where the technique was pioneered. Surgical teams took immature fetal cells destined to become dopaminergic neurons from the midbrain of aborted fetuses and transplanted them into the striatum of patients' brains, the area of greatest dopamine loss in Parkinson's disease.

More than 100 patients worldwide received the therapy as part of clinical trials before the moratorium. "But centres used different procedures and protocols — it was impossible to work out why some patients did very well and others didn't benefit at all," says Barker.

In 2006, Barker, together with neuroscientist Anders Björklund at Lund University, set up a network to bring together the original seven teams that had performed the transplants, to assess all protocol details and patient data retrospectively.

The teams worked out that the procedure tended to be most effective in patients who were relatively young and whose disease was at an early stage. In addition, post-mortem analysis of patients' brains showed that those who benefited most had at least 100,000 dopamine-producing cells of fetal origin integrated into their brains. Cells from at least three fetuses are needed to achieve these numbers, the neuroscientists concluded.

The retrospective analysis encouraged the European scientists, including Barker and ►



► Björklund, to launch a new trial, which is funded by the European Union, involving fetal dopaminergic-neuron transplants. Known as TRANSEURO, it will monitor disease progression in 150 patients in the United Kingdom, Sweden, France and Germany. The first patient is due for transplantation next month at Addenbrooke's Hospital in Cambridge. In line with the retrospective findings, the average age of trial participants at recruitment was 55, and their average disease duration just 4 years. None had displayed dyskinesias — uncontrolled muscle movements that can be a side effect of L-dopa treatment.

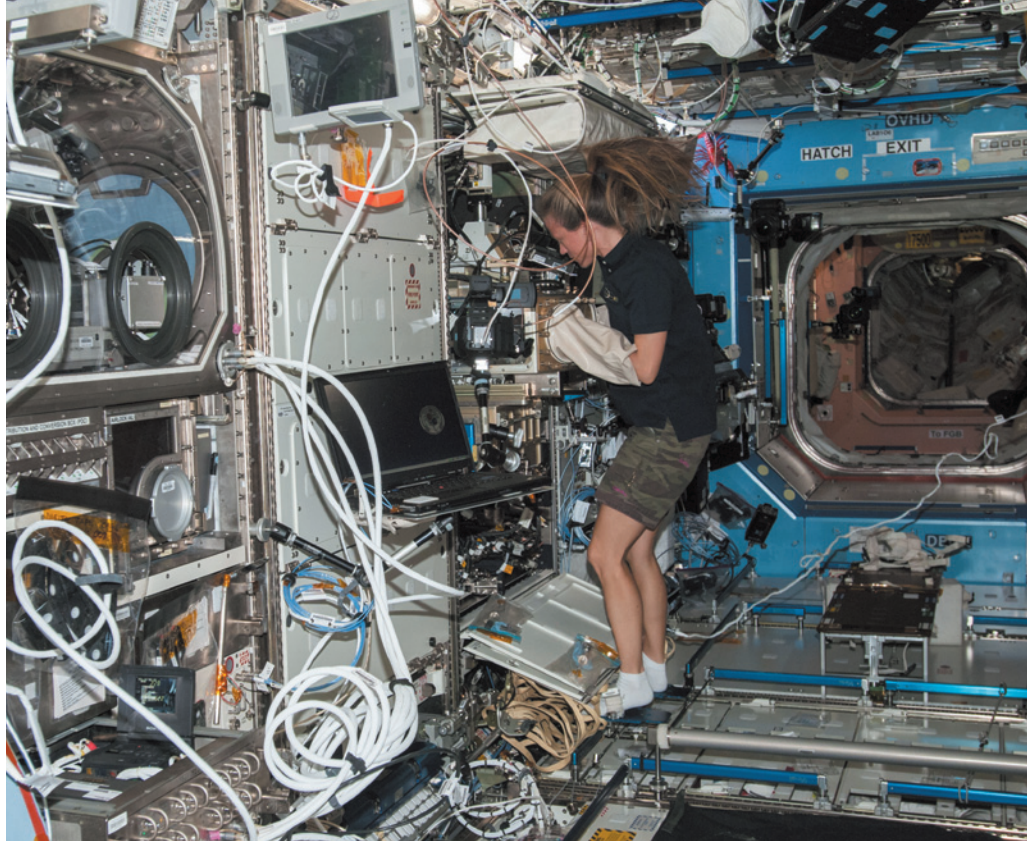
But stem-cell biology has advanced significantly since 2003, and dopaminergic neurons can now be derived from human embryonic stem cells and also from induced pluripotent stem cells — mature cells that have been rewound to an uncommitted stem-cell-like state and that can be coaxed to become a cell type of choice. These potential sources are more desirable than those derived from fetuses, because fetal cells are hard to come by and their biology varies.

Research is under way to ensure that the stem cells develop into the exact type of dopaminergic cell needed to treat Parkinson's and that they become correctly integrated into recipients' brains. But progress has been so fast that clinical trials are already on the horizon. A Japanese trial, using induced pluripotent stem cells, is planned to start in Kyoto within two years; and two trials using human embryonic stem cells are also planned, one to begin within three years in New York and the other in Europe within four to five years.

The Parkinson's Disease Global Force hopes that its joint planning will make comparing outcomes easier. Members will share their protocols for deriving and grafting cells, as well as their clinical criteria for patient selection and follow-up.

They see the TRANSEURO trial as a pathfinder. "We don't know yet which source of cell will turn out to be the best, but right now the fetal cell is the gold standard we need to match," says neurologist Claire Henchcliffe from the Weill Cornell Medical Center in New York, who is coordinating the working group's guidelines on patient assessment and trial design.

The stem-cell approaches have a long way to go before they can rival the promise of fetal cells, says Lund University stem-cell biologist Malin Parmar, a member of the European clinical-trial team. That is because the cells from fetal brains are already on the way to becoming mature dopaminergic cells. "The human body knows very well how to develop each cell type from the embryo," she says. "We haven't learnt all of these secrets yet, but we have learnt some major ones." ■



NASA astronaut Karen Nyberg works on a colloid experiment aboard the International Space Station.

#### SPACE

# Space-station science ramps up

*NASA pushes research agenda in face of Russian resistance.*

BY ALEXANDRA WITZE

In January, when the United States proposed extending International Space Station (ISS) operations until 2024, the world was a very different place. That was before Russian military intervention in Ukraine, before US–Russian relations foundered and before Russian deputy prime minister Dmitry Rogozin suggested that US astronauts use a trampoline to get themselves to orbit (see *Nature* <http://doi.org/s4f>; 2014).

Rogozin also suggested last month that Russia would stop participating in the space-station programme after its scheduled end date of 2020. That statement did not set official government policy, but given Russia's key role in the orbiting outpost it cast a shadow over hopes for the four-year extension.

With the clock ticking, the race is on to conduct as much science as possible in whatever time the space station has left. At a conference next week in Chicago, Illinois, NASA scientists will try to lure researchers who have not worked with near-zero-gravity conditions before. The goal is to get them to

propose anything from the usual research agenda — such as protein crystallization and human physiology experiments — to basic biomedical research and Earth-science observations that can take advantage of the high-flying platform before its mission ends (see 'Research push'). "There's never been anything like it," says Julie Robinson, NASA's space-station research chief at the Johnson Space Center in Houston, Texas. "It's like a university, all together with all the disciplines — I don't know if we'll see that again."

More than 1,600 scientists from 69 countries have contributed to experiments carried out on the space station since its first module was launched in 1998. The United States is the largest science user. But over the years, many have questioned the value of the science done in orbit. One main goal is to help humans to endure long-duration spaceflight, but early experiments often failed. For instance, NASA astronauts would spend hours a day exercising on treadmills to slow down muscle wasting and bone loss — to little avail. Force measurements revealed that they were subjecting their bodies to stresses that were not

NASA



## RESEARCH PUSH



The United States has proposed extending the International Space Station's lifetime by four years, to 2024, to conduct more scientific research. This year, the station will see the arrival of multiple science payloads.

**1 JULY** Launch of Orbital Sciences cargo ship, which includes student experiments.

**25 JULY** Launch of Automated Transfer Vehicle cargo ship and experiment supplies.

**AUGUST** Launch of SpaceX ship, including rodent habitat and RapidScat ocean-wind monitor.

**12 SEPTEMBER** Launch of SpaceX cargo ship with cloud-aerosol laser instrument.

even close to the pull of gravity on Earth.

Another goal has been to conduct basic scientific observations to see how physical and biological processes change in a near-weightless environment. But these studies have often been limited to growing relatively unimportant proteins or running student experiments, and have often not made fundamental breakthroughs. In a sharply critical 2011 report, a US National Research Council panel argued that NASA was “poorly positioned to take full advantage of the scientific opportunities offered by the now fully equipped and staffed ISS laboratory”.

Space-station research suffered further after NASA halted the space-shuttle programme in 2011, eliminating the only option

for both delivering experiments to orbit and returning them to Earth.

But in late 2012, the debut of two-way cargo ships made by private company SpaceX of Hawthorne, California, again allowed samples to be returned to Earth. That development enabled experiments such as those of David Klaus, an aerospace engineer and gravitational biology researcher at the University of Colorado Boulder, who is exploring why antibiotics seem to work less well in space than on the ground. The answers might help in making the drugs perform better on Earth, he says — and running operations to 2024 would allow him more generations of experiments, with the results of each guiding the design of the next.

“Extending the station out that long helps us align a bit better with research portfolios here on the ground,” says Duane Ratliff, chief operating officer at the Center for the Advancement of Science in Space, an organization in Melbourne, Florida, that manages US space-station research for NASA.

Another strategy is to mount Earth-science experiments on the space station as a cheaper alternative to putting them on a free-flying satellite. Some time in August or later, an instrument that monitors ocean winds will be flown to the station to replace a satellite-borne one that failed in 2009. And a supply run in September is slated to deliver a laser system to measure clouds, dust and pollution in the Earth's atmosphere.

The agency is also adding what it hopes will be versatile facilities that produce data for a broad range of research. The August supply run will carry a rodent habitat — the largest ever launched for long-duration spaceflight, says Robinson, with a capacity of 40 mice. And a series of experiments called geneLAB will send a range of model organisms, including fruit flies and nematodes, into space for months at a time, performing basic biomedical assays on them both in orbit and after returning them to Earth. The accumulated information will go into a database that any ground-based researcher will be able to draw on.

“People will come and do one to two experiments in space,” says Robinson, “and continue to do work in their lab for another 30 years to understand that insight.” ■



**MORE  
ONLINE**

## VIDEO OF THE WEEK



Artists turn simulated ecosystem into symphonic forest soundscape  
[go.nature.com/oio1cn](http://go.nature.com/oio1cn)

## MORE NEWS

- Head of Science Europe talks to *Nature* [go.nature.com/5ey2tv](http://go.nature.com/5ey2tv)
- Nuclear waste leak shuts down neutrino experiment [go.nature.com/ziemv9](http://go.nature.com/ziemv9)
- Rock isotopes confirm Moon-formation theory [go.nature.com/oknuzq](http://go.nature.com/oknuzq)



Demonstrators protest against gene patenting outside the US Supreme Court last April.

## GENETICS

# Cancer-gene data sharing boosted

*Efforts to get more breast-cancer gene variants into public databases are gaining ground.*

BY ERIKA CHECK HAYDEN

When the US diagnostics giant Myriad Genetics had its legal monopoly on breast-cancer gene testing eliminated one year ago, the company still retained an enormous edge over competitors. Although the US Supreme Court's ruling last June invalidated the patenting of genes, and with it Myriad's exclusive rights on two genes associated with breast- and ovarian-cancer risk, the firm still has a private trove of data from 1.3 million genetic tests.

That information gives Myriad, of Salt Lake City, Utah, an advantage in interpreting test results on these genes.

But a coalition of scientists, physicians, patients and genetic counsellors says that it will soon eliminate that advantage. A year after the Supreme Court invalidated the patenting of genes — and with it, Myriad's monopoly on testing for mutations in the *BRCA1* and *BRCA2* genes linked to breast and ovarian cancer — the number of entries for *BRCA* variants in ClinVar, a public database for clinical genetic data, has grown to around one-third of the number in the

Myriad database. Leaders of the public effort say that it is a showcase for how scientists can clear long-standing obstacles to sharing genetic information.

"My guess is that within a year's time, we would have a more robust data source for *BRCA1* and *BRCA2* than Myriad has by itself today," says geneticist Heidi Rehm, head of the Laboratory for Molecular Medicine at the non-profit company Partners HealthCare in Cambridge, Massachusetts, and a central figure in the *BRCA* data-sharing drive.

## PUBLIC VERSUS PRIVATE

Myriad counters that public databases are unreliable, expensive and vulnerable to funding cuts that compromise their upkeep. "We have the highest-quality databases in the world," says company spokesman Ronald Rogers. "And that's important because when the patient is given a result, they're going to make a medical management decision based on that information." Rogers says that the firm spent US\$500 million to develop its tests and database.

In genetic testing, the size of the reference database matters: the bigger it is, the more

useful it becomes for interpreting the results of any individual gene test. Of the thousands of possible variations on the spelling of the DNA bases that make up the *BRCA* genes, only some will be linked to cancer. If a particular variant is significantly more frequent among those who develop cancer than in the general population, its contribution to disease risk can be calculated. Only by collecting data from many different people can scientists observe the same variants often enough to make these kinds of calculations with confidence.

"If Myriad holds the data hostage in a proprietary database, they're harming patients," says Sherri Bale, managing director of GeneDx in Gaithersburg, Maryland, one of several competitors being sued by Myriad for infringement of other patents associated with the *BRCA* tests. These companies have cut the price of *BRCA* tests to an average of \$2,200 compared with Myriad's \$4,040 — and are contributing an accelerating flow of *BRCA* data to ClinVar, which is held at the National Center for Biotechnology Information in Bethesda, Maryland.

"All of these companies seem much more willing to share data," says genetic epidemiologist David Goldgar of the University of Utah's Huntsman Cancer Institute in Salt Lake City.

Geneticists such as Rehm and Robert Nussbaum at the University of California, San Francisco, are urging patients, genetics professionals and insurance companies to use only *BRCA*-testing companies that share data.

Geneticists have long exhorted their colleagues to share, but have been blocked by practical and competitive barriers. Sharing takes time and money, and geneticists fear compromising patient privacy or seeing competitors beat them to market or publication with results based on their own data. An initiative started by Nussbaum — the Sharing Clinical Reports Project — aims to bring about a sea change by coaxing clinical labs to deposit the results of Myriad diagnostic tests into ClinVar.

ClinVar now holds information on 5,752 *BRCA* variants, deposited by Sharing Clinical Reports, Myriad competitors and academic labs, compared with Myriad's 16,000. And other, broader, plans for data sharing are afoot. At a meeting in London in March, geneticists and physicians convened by the non-profit Global Alliance for Genomics and Health, based in Toronto, Canada, which promotes data sharing, outlined a plan to create an even more extensive data resource. Dubbed the *BRCA* Challenge, it would link major clinical-genetics databases such as ClinVar and the Leiden Open Variation Database, funded by the European Union and run by a team in the Netherlands.

"The thought," Goldgar says, "is that the rest of the world has roughly the same amount of data as Myriad — and if we put it all together and make a concerted effort to try to use it, then we can be on equal footing." ■

TOM WILLIAMS/CQ ROLL CALL/GETTY



## ACOUSTICS

# Sound clue in hunt for MH370

*Hydroacoustic signal caught by sensors in the Indian Ocean may be linked to crash of Malaysian airliner.*

BY DECLAN BUTLER

Researchers are exploring what may be the first promising lead in months in the search for the wreckage of Malaysian Airlines flight MH370. Sensitive microphones on the ocean floor off Australia picked up a distinctive signal at about the time that the Boeing 777 aeroplane is believed to have crashed in the Indian Ocean. The findings, announced by Australian scientists last week, offer a rough location for the source of the sound and are being followed up by search authorities.

The signal was discovered by a team led by Alec Duncan, an underwater acoustics specialist at Curtin University's Centre for Marine Science and Technology in Perth, Australia. It was recorded at 01:30 coordinated universal time (UTC) on 8 March; the last transmission from flight MH370, an electronic 'handshake' between the aircraft's communication system and a satellite, took place at 00:19 UTC, when the plane is estimated to have run out of fuel.

The sound is believed to have originated somewhere along a strip running to the northwest of the Indian Ocean (see 'Sound trackers'). That is out of the range of the current search, which was determined by analysis of the satellite communication data and is being led by the Australian Transport Safety Bureau (ATSB). However, the techniques used are well-established. "The ATSB will continue to discuss the analysis of this information with Curtin University for the purposes of informing the search," says a spokesperson for the Joint Agency Coordination Centre (JACC) in Canberra, which is coordinating the Australian government's support for the search.

The Curtin team emphasize that the sound may have come from other sources, such as a small earthquake, but thinks that the lead is worth pursuing. It is now preparing to retrieve more hydroacoustic data from the ocean off northwestern Australia.

Duncan's team found the signal while analysing data from an acoustic station in Perth Canyon about 40 kilometres west of Rottnest Island near Perth. It is one of six stations operated by Australia's Integrated Marine Observing System (IMOS), which was set up to make physical, chemical and biological observations of the ocean basin. Duncan then confirmed the signal using data from an acoustic station off Cape



Hydrophones run by the Comprehensive Nuclear-Test-Ban Treaty Organization track explosions in the sea.

Leeuwin on the southwest tip of Australia that were provided by the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO).

The CTBTO, a nuclear-test monitoring body based in Vienna, maintains a global network of seismic and radioisotope detectors, as well as other instruments. The network includes six hydrophone stations that monitor for explosions in the ocean, but can also pick up other sounds such as whale calls at great distances. Cape Leeuwin is one of two CTBTO acoustic stations in the Indian Ocean. Duncan also analysed data from the other, off Diego Garcia island in the middle of the Indian Ocean, but found

**"Underwater acoustic data still has the possibility of adding something to the search."**

nothing. The data were too polluted by noises from seismic surveys, he says.

The IMOS stations have just one microphone each, so alone cannot provide detailed information on the direction of sounds. But the CTBTO's stations have two sets of three hydrophones separated by several kilometres, which — like a pair of human ears — allow listeners to get a fix on a sound's direction to within 0.5°.

Duncan and his colleagues now plan to recover and analyse data from the two IMOS stations off northwestern Australia. But those loggers record only 5 minutes of sound every 15 minutes, and any signals are likely to be contaminated with noise from seismic surveys, says Duncan. He reckons that there is only a "slight chance" that their data will contain the signal, but that it is "worth a go". The team had planned to recover the sensors in September or



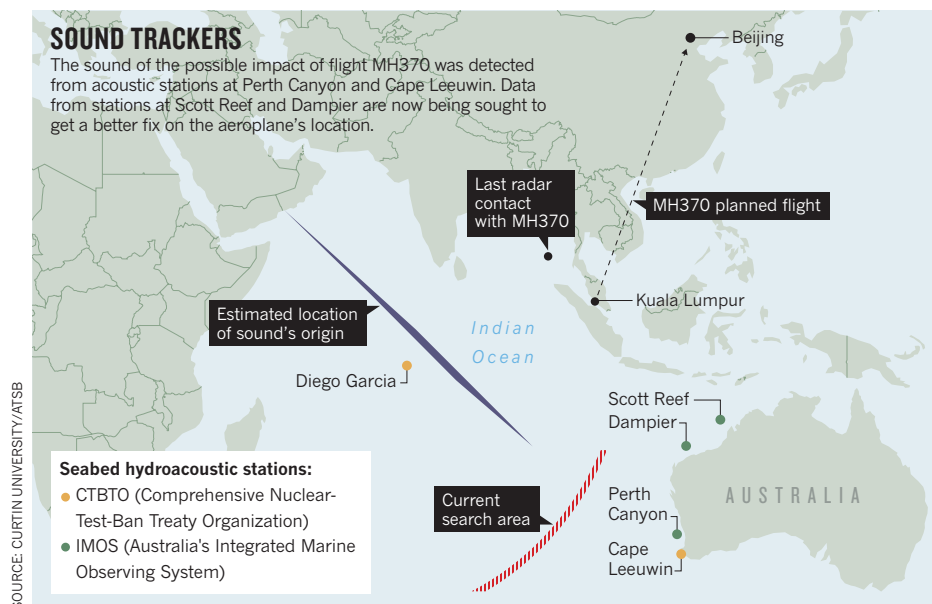
October, but now hopes to make the 7–9-day round trip in August. “Given the continuing uncertainty regarding the fate of MH370, underwater acoustic data still has the possibility of adding something to the search,” says Mark Prior, a CTBTO seismic-acoustic officer.

Meanwhile, it is unclear what other sources of hydrophone data that could be used in

the search exist in the region. The US Navy deployed vast arrays of hydrophones on the ocean floor during the cold war for anti-submarine warfare. Details of the Sound Surveillance System (SOSUS) remain secret, but most of the hydrophones are thought to have been deployed off the US Atlantic and Pacific coasts.

Once the cold war ended, however, the system was downgraded. The data are still being collected, but they are not routinely analysed unless there is an underwater threat. William Marks, a spokesman for the US Navy in Yokosuka, Japan, declined to comment on whether the United States had hydrophones in the region. “Discussions of the SOSUS system at that level are classified,” he says. “This is a very sensitive system.”

India and Pakistan also have submarine fleets, but Duncan and other scientists say that they do not know whether they or any other nation has hydrophones in the Indian Ocean. “We have not been advised of any hydrophone facilities operated by India or Pakistan,” says the JACC spokesperson. ■



### CORRECTIONS

The News story ‘Phage therapy gets revitalized’ (*Nature* **510**, 15–16; 2014) mischaracterized the CRISPR mechanism for tackling antibiotic-resistant microbes. It should have said that the phage injects DNA into the bacterium, which then transcribes it into RNA. And in the News story ‘Chicken project gets off the ground’ (*Nature* **509**, 546; 2014), the mentions of ‘guinea fowl’ should have read ‘jungle fowl’.

# STARS OF SOUTH AMERICAN SCIENCE

*Growing resources for research and development are creating opportunities across the continent, but many countries still struggle to build their programmes.*



**SPECIAL  
ISSUE**

**L**ike the night sky, the overall sweep of science in South America can look pretty dark. Brazil is the only country on the continent that spends more than 1% of its gross domestic product on research and development, and even that investment sits far below what other countries of similar means are ploughing into science.

But take a closer look at the continent's scientific enterprise, and bright spots emerge. At the start of the FIFA World Cup in Brazil, with billions of people focusing on South America, *Nature* examines a part of the world that has spent too long on the sidelines of science.

A graphic tour on page 202 details the inputs and products of research and development on the continent. The region faces many challenges in terms of building a strong scientific workforce and boosting resources, but investment and publications are climbing. A News Feature on page 204 profiles several key institutions and research groups — from agricultural specialists in Colombia to RNA experts in Argentina — who have gained worldwide recognition.

An Editorial on page 188 calls on international colleagues to help build South American science in ways that do not cause young researchers to leave permanently. On page 213, a Comment describes one such success: the Pew Latin American Fellows Program, which each year sends about ten top graduates to work in North American labs. More than 70% return to their native countries, bringing with them the expertise they have gained. That initiative is a part of broader efforts, described in a News Feature on page 207 that examines how countries are trying to repatriate scientists who left to train abroad.

As economies on the continent heat up, they are devoting greater resources to research, increasing the need for better infrastructure and policies to support science. In a Comment on page 209, research leaders describe how they hope to navigate this growth, and how science can help to expand their countries' economies sustainably. Ideas range from creating a science ministry to using research to find new commercial uses for the fruits of the Amazon.

Many researchers in South America maintain a cautious outlook — they have lived through periods of intense economic and political strife in the not-too-distant past. But they also harbour the hope that the continent's science is headed for a winning season. ■



**SOUTH AMERICAN SCIENCE**

A *Nature* special issue  
[nature.com/southamerica](http://nature.com/southamerica)

# SOUTH AMERICA

## by the numbers

By Richard Van Noorden

The expanding economies of South America have led to a significant rise in scientific output over the past two decades, and research spending has increased in most countries. But given the region's share of the world's population and gross domestic product (GDP), publication rates still fall short of what would be expected. Research quality has not kept

pace with rising output, and the continent's research papers still struggle to attract citations from the rest of the world. There are huge inequalities across the region, too: Brazil dominates the publication record, for example, whereas Chile takes pole position in the patent landscape and Argentina scores highly in terms of the proportion of its population working in science.

### THE PUBLISHING LANDSCAPE

South America has boosted its share of the world's research articles — but at 4%, it still underperforms slightly relative to its 5–6% share of world population and GDP.

NUMBER OF ARTICLES PUBLISHED IN ELSEVIER'S CITATION DATABASE SCOPUS IN 2013 (see 'The hidden continent' below)

#### BRAZIL: 46,306

In the past 20 years, Brazil's scientific output has risen by more than a factor of five, as its economy has almost tripled in terms of purchasing power. The country now accounts for more than two-thirds of South America's entire research output — although it is broadly similar to Argentina, Uruguay and Chile in terms of articles per capita.

#### VENEZUELA: 1,315

The only South American nation whose scientific output is declining: its publication tally fell by 29% between 2009 and 2013.

#### PERU: 1,044

Nearly three-quarters of Peru's articles involve collaborations with other countries. The most-cited articles include work on prevention of HIV, tuberculosis and lupus.

#### PARAGUAY: 87

#### ARGENTINA: 9,337

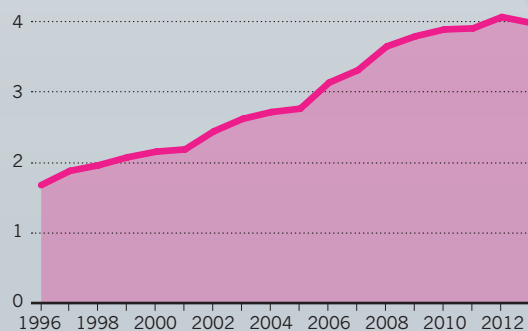
Has hauled up the impact of its research to just above the world's average — outperforming Brazil.

#### URUGUAY: 799

#### CHILE: 6,794

As well as its astronomical observatories, the country has also found scientific success working on food crops, such as a highly cited collaboration on the genome of the potato.

South American share of world publications (%)



### The hidden continent

South America's research strength may be underestimated because its researchers often publish in journals that are not indexed in major citation databases, such as Elsevier's Scopus or Thomson Reuter's Science Citation Index. In 2012, for example, some 6,000 of the roughly 20,000

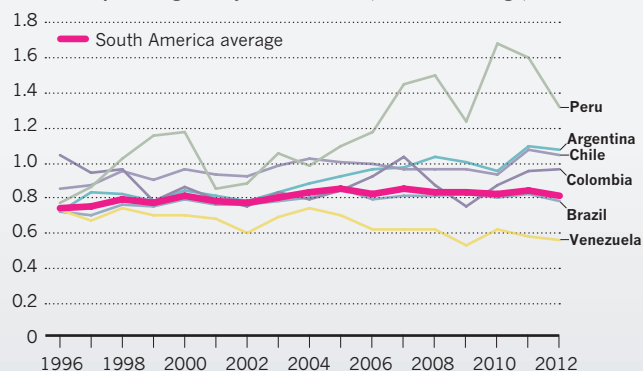
papers that Brazil published in SciELO (Scientific Electronic Library Online), a subsidized collection of mainly Latin American journals, were not indexed in Thomson Reuter's database. But last year, Thomson Reuters agreed to create a database for the SciELO index.



## COLLABORATION AND EXCELLENCE

South America's scholarly impact remains relatively low — its citation rate last year was around 80% of the world's average (below). Peru's articles do best, largely because most are co-authored with scientists outside the continent. Indeed, the region's less-developed countries are generally more likely to collaborate beyond South America. In Brazil, less than one-quarter of its articles in 2008–12 involved such partnerships (right).

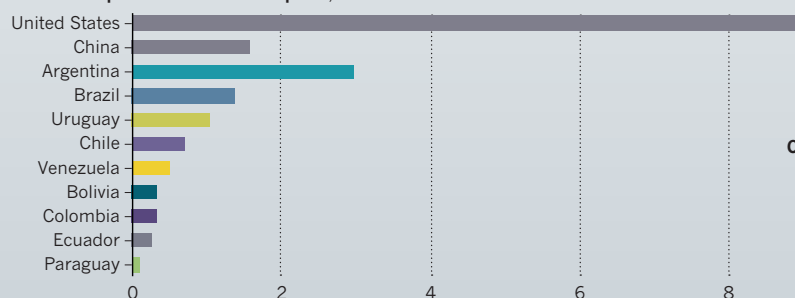
Citation impact weighted by research field (1 = world average)



## RESEARCH STRENGTH

Brazil has more than 100,000 full-time researchers, single-handedly providing nearly two-thirds of South America's science personnel. But Argentina has the greatest proportion of researchers, with almost 3 scientists for every 1,000 workers.

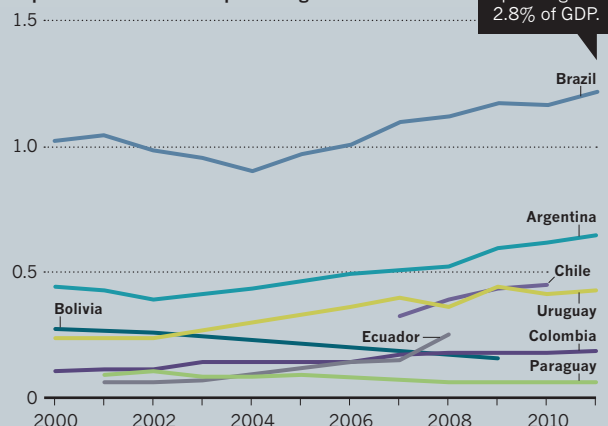
Full-time equivalent researchers per 1,000 labour force



## RESEARCH SPENDING

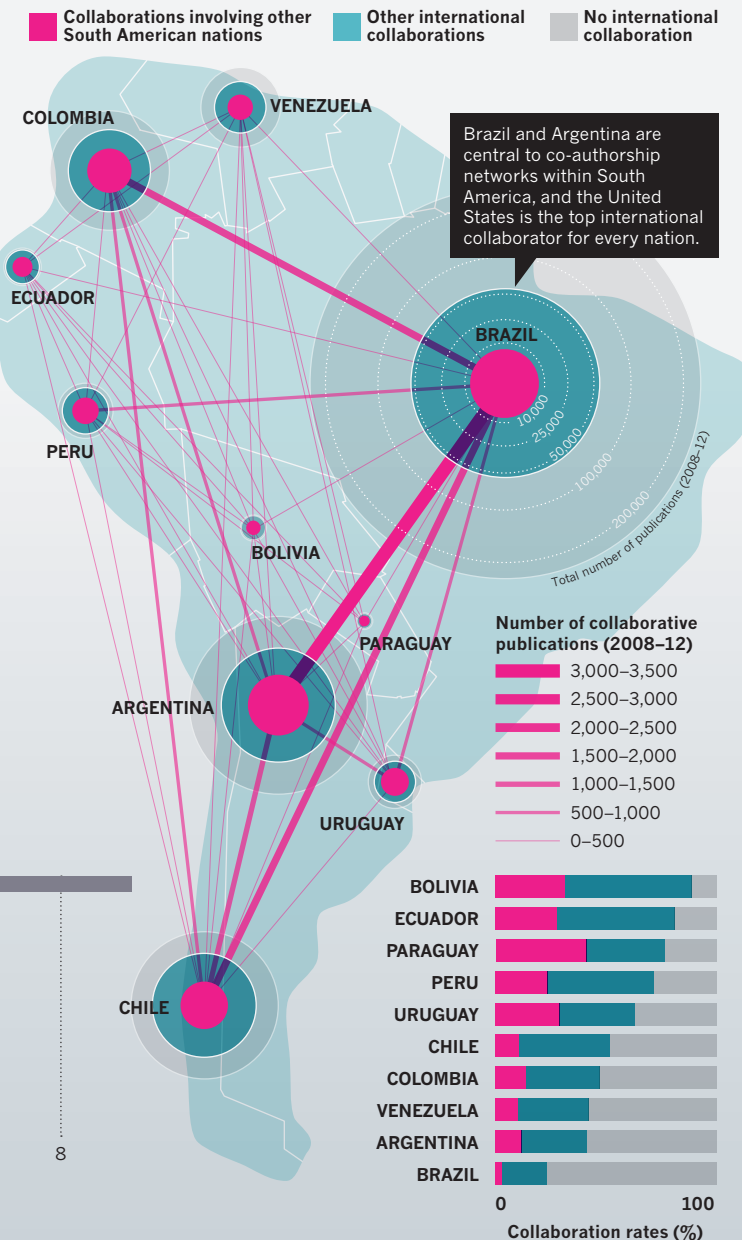
Argentina and Brazil's spending on research and development (R&D) has shot up even faster than their economies have grown. Brazil remains the region's only country to devote more than 1% of its economy to R&D\*.

Expenditure on R&D as a percentage of GDP



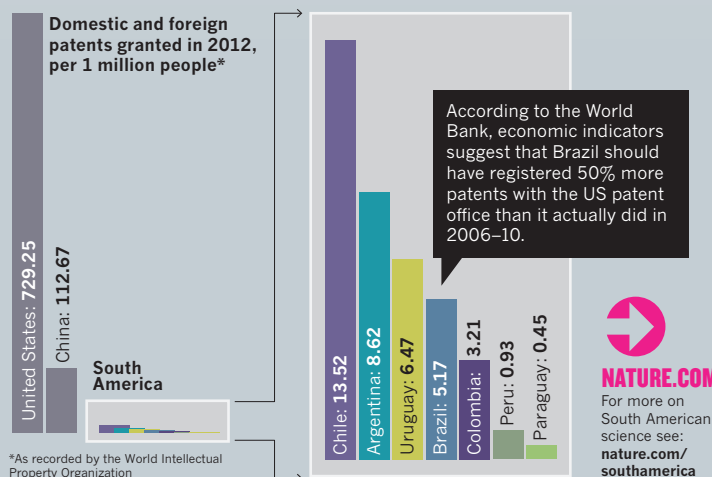
\*No verified figures for Venezuela, no up-to-date data for Peru. Data are incomplete for Ecuador and Chile.

In 2011, US spending was 2.8% of GDP.



## PATENTS GRANTED

In Brazil, nearly half of research funding comes from the business sector; in other South American nations, the share from businesses is generally much lower, a stark contrast with many industrialized countries. Poor private investment results in a small number of patents granted per capita, where South American countries look particularly weak.



**NATURE.COM**  
For more on South American science see: [nature.com/southamerica](http://nature.com/southamerica)

# Big *players*

Despite myriad problems in many countries, pockets of excellence thrive in South American science.

**IT MAY SEEM HERETICAL TO SAY SO IN THE LAND OF THE BEAUTIFUL GAME,** but science in Brazil beats the World Cup — at least in a financial match-up. Government and businesses there invest some US\$27 billion annually in science, technology and innovation, dwarfing the price tag for the football tournament, which tops out at about \$15 billion.

Science in Brazil and many other countries in South America has come a long way since the dark days of the dictatorships just a generation ago. In Argentina, the number of science doctorates jumped nearly tenfold between 2000 and 2010; Peruvian scientists tripled the tally of articles they produced over the same period; and science funding is climbing in most countries.

South American science still has far to go if it hopes to catch up with other continents. By many measures — such as investments, patents and education — the countries there lag behind other nations with similar levels of gross domestic product (GDP). There is looming instability in countries such as Argentina and Brazil, where recent protests reflect deep social and economic divisions — problems that plague much of South America. But amid the concerns, there are many bright spots in the world of science. Here, *Nature* highlights several examples of outstanding researchers and institutions in the region.



CHILE

## UPWARD TRAJECTORY

BY MICHELE CATANZARO

**W**hen Mario Hamuy finished his university degree in Chile in 1982, he was the only one in the country interested in pursuing graduate studies in astronomy. Now, more than 25 Chilean students join such programmes each year and Hamuy directs the Millennium Institute of Astrophysics in Santiago, home to 95 students and faculty members.

During the course of Hamuy's career, Chile has emerged as a major player in the world of international astronomy, in no small part because of the extraordinary collection of telescopes housed in the country's highlands.

**The European Southern Observatory operates the Very Large Telescope in northern Chile.**

"Astrophysics has come to the forefront of Chilean science thanks to the increase in human resources and to the fact that

B. TAFRESH/ESO





## BRAZIL

## SÃO PAULO'S HEAVY HITTER

BY GIULIANA MIRANDA

**A**lthough Brazil rivals Europe in size, much of the leading research in South America's largest country emanates from an area the size of the United Kingdom. São Paulo, in southern Brazil, is the richest of the country's 26 states and publishes more than half of Brazil's scientific articles. One of the main reasons

we have the cleanest sky in the world," says Dante Minniti, an astronomer at the Pontifical Catholic University of Chile in Santiago.

Although Chile invested just 0.44% of its GDP in scientific research in 2011, the latest year for which figures are available, funding for astrophysics has steadily grown, from \$2 million in 2006 to \$6.8 million in 2010. Over the same period, the number of faculty positions has almost doubled. And the country's publications in astronomy have risen more than fourfold during the past decade.

The quality of the work has improved as well. Chile ranks highly in terms of citations per paper in space science, and some of its scientists have made important discoveries. In the early 1990s, Hamuy made a key contribution that helped others to measure the accelerating expansion of the Universe and win a Nobel Prize in 2011. And Minniti is one of the leaders at the VISTA infrared survey telescope at the European Space Organization's Paranal Observatory in northern Chile, which has created a catalogue of more than 84 million stars in the central parts of the Milky Way.

Chile's skies have been attracting international telescopes since 1964. By 2020, when the European Extremely Large Telescope is due to be completed, the country is expected to host 70% of the global observation surface for large optical and infrared telescopes.

By contract, Chilean astronomers get 10% of the observation time on each telescope

for its success is the São Paulo Research Foundation (FAPESP), the state agency that promotes research and education. In 2013, the agency invested \$512 million in science funding, more than many nations in the region. (At the federal level, Brazil's National Council for Scientific and Technological Development has a budget of about \$650 million for science, technology, and innovation in 2014.)

installed in the country. But some astronomers say that this is too little, considering how much the country provides for the organizations running the telescopes.

"This country has given enormous advantages to the international consortia, ranging from full tax exemption to diplomatic status: it's time that Chile participates in a more active way," says Mónica Rubio, director of the astronomy programme of the Chilean funding agency CONICYT.

A unanimous aspiration of Chilean scientists, says Rubio, is not just to use observatories but also to build them, through local companies and engineers. Another plan Rubio is working on is developing the Atacama Astronomical Park, a 36,347-hectare protected area around the Atacama Large Millimeter/submillimeter Array, which CONICYT plans to use to attract future telescopes from Brazil and the United States, and maybe also from China, South Korea and Thailand.

But many astronomers are worried about the governance of science in Chile. CONICYT has lacked a director since José Miguel Aguilera resigned eight months ago, and the country's new president, Michelle Bachelet, has frozen plans to create a science ministry (see *Nature* 507, 412–413; 2014). "It's a good moment for Chilean astronomy, but keeping the momentum will require more sustained support from the government," says Minniti. ■

**A ZnO semiconductor from a FAPESP-funded project.**

Created in 1960, FAPESP has a stream of funding guaranteed by the constitution of São Paulo, which requires that 1% of the tax revenue goes to the foundation. Its success in fostering research and education inspired other Brazilian states: all but one now has a similar agency, and most have guaranteed funding linked to taxes.

FAPESP directs 37% of its funding to basic research in fields ranging from climate change to particle physics. About 10% goes to infrastructure and the rest is channelled to applied research. Nearly one-third of its total budget is devoted to medical research.

"One difference in FAPESP's work is that we invest a lot in basic science," says Carlos Henrique de Brito Cruz, FAPESP's scientific director. "We believe in balance."

The most recent large project approved for funding is the Long Latin American Millimeter Array radio telescope, a joint project between Brazil and Argentina that will receive \$12.6 million from the agency and an equal amount from Brazil's science ministry. FAPESP's board is considering a \$40-million investment in the Giant Magellan Telescope, which would give São Paulo astronomers access to the facility, planned for construction in Chile.

Science officials in other nations can only look with envy at the agency's guaranteed funding. "FAPESP is a very interesting model for us because São Paulo is one of the few states in the world where support of research is linked directly to GDP," says Martyn Poliakoff, foreign secretary and vice-president of the Royal Society in London.

Regional agencies such as FAPESP play a very important role in Brazil, says Wanderley de Souza, a biomedical scientist at the Federal University of Rio de Janeiro and a member of the Brazilian Academy of Science. "They can make research happen even if the federal funding gets scarce."

Brazil struggles with vast economic differences among its various regions, and that is reflected in regional science budgets. FAPESP has the biggest budget of all the regional agencies, but that does not reduce federal investments in the state, says Clelio Campolina, the minister of science, technology and innovation. "We want to improve other states, but also reward excellence," he says.

FAPESP's rapid growth has raised some concerns among scientists in São Paulo who complain about an increase in bureaucracy. But agency officials defend its performance and say they are working to improve its procedures.

It's all part of an effort to produce high-quality work, says Brito Cruz. "We want the best projects." ■



SOUTH AMERICAN SCIENCE

A *Nature* special issue  
[nature.com/southamerica](http://nature.com/southamerica)



# COLOMBIA GROWTH CENTRE

BY LISA PALMER

In the Cauca Valley of western Colombia, a herd of hefty cows at Petequi farm munches away on lush grass that looks as if it has grown there forever. But the plants are relative newcomers. They are cultivars of African super grasses, bred for enhanced nutrition and hardiness by researchers at the International Center for Tropical Agriculture (CIAT), less than 50 kilometres to the north.

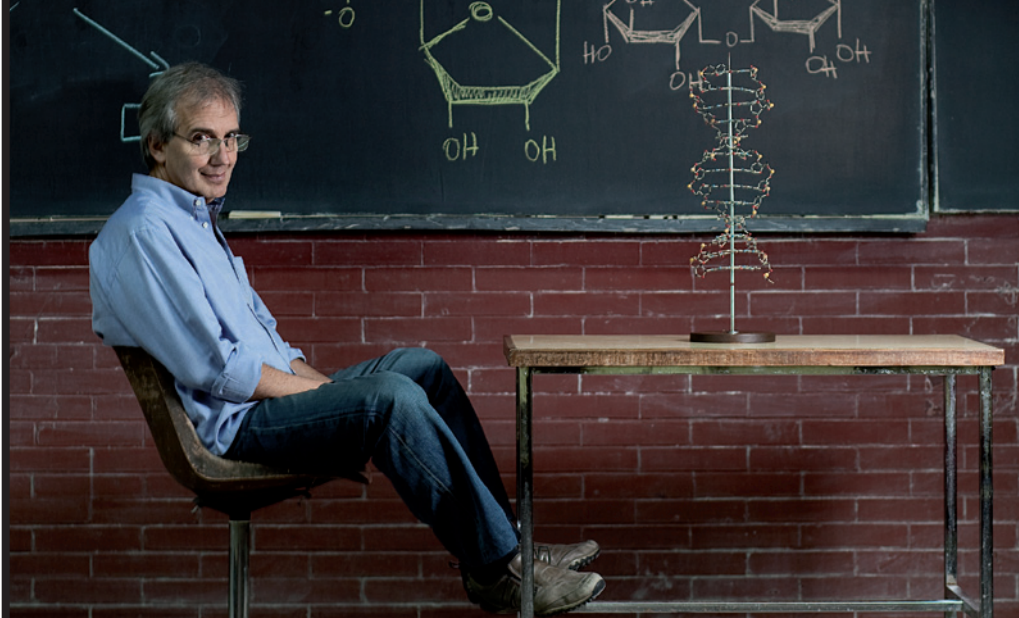
Cows at Petequi once took four years to reach market weight. Now they fatten up in just 18 months. The story is much the same throughout the South American *cerrado*, or savanna. The improved grasses have revolutionized tropical forage across the continent thanks to the combined work of researchers at CIAT and the Brazilian Enterprise for Agriculture Research, a state-owned Brazilian company, says Eduardo Trigo, an agricultural economist and science adviser to the Argentine ministry of science, technology and innovation in Buenos Aires. "CIAT has been one of the key actors in the development of the South American *cerrado*," he says.

Established in 1967, the Colombian facility was one of the first members of the CGIAR consortium of international agriculture research centres. CIAT employs 325 scientists and has an annual budget of \$114.4 million, paid for by the multi-donor CGIAR fund and by other international donors.

Aside from its work on grasses, CIAT has focused on breeding improved varieties of beans, rice and cassava — staple crops that are important to the food security of the rural poor. "Genetic improvement of these crops has proved to be a powerful weapon for combating hunger and poverty," says Ruben Echeverria, director-general of CIAT. For example, beans developed by CIAT from Latin American varieties are now feeding up to 30 million people in Africa, according to the centre.

Some 70% of rice in South America, and 90% of cassava in Asia, can be traced back to CIAT's breeding programme. "Cassava is now a multibillion-dollar business for starch production in Asia, providing income to smallholders," says Andy Jarvis, leader in policy research at CIAT.

The centre has also helped to grow expertise on the continent and elsewhere; since CIAT opened, some 13,000 researchers have trained there. Its facilities have been instrumental in building capacity for plant physiologists in the poorer countries of the Andean region, says Trigo. ■



Alberto Kornblihtt:  
RNA pioneer.

## ARGENTINA THE RNA SLEUTHS

BY ALESZU BAJAK

Molecular biologist Alberto Kornblihtt likes to put things in perspective. "We may be on the periphery of scientific research," he admits from his office in Buenos Aires. "But it's not an impossible place to do science." In fact, he and his community of researchers in alternative RNA splicing — a field he helped to create — have shown that they can do world-class research despite tight government budgets and three-month delivery times for reagents that can cost three times as much as they would in the United States or Europe.

Like Kornblihtt's lab, alternative RNA splicing makes use of constrained resources in innovative ways. Through varied patterns of cutting and rejoining, a single transcribed gene can give rise to many different messenger RNAs, thus permitting a single gene to express different proteins. Kornblihtt found one of the first cases of this process in humans while he was a postdoctoral fellow in the United Kingdom. He moved back to Argentina in 1984 and has assembled a group of researchers that continues to explore this realm.

It has been a good year for his group. Kornblihtt and his doctoral student Ezequiel Petrillo published a paper in *Science* in April on how light affects alternative splicing in plants (E. Petrillo *et al. Science* <http://doi.org/s2d>; 2014). And last month, Gwendal Dujardin, a postdoctoral fellow from France (a rare sight in an Argentine lab), published a splicing study in *Molecular Cell* (G. Dujardin *et al. Mol. Cell* **54**, 683–690; 2014).

The work is all part of a continuum, says Kornblihtt. He considers scientific research in his native Argentina to be part of a long tradition that started with Bernardo Houssay and Luis Leloir, twentieth-century Nobel laureates whose names now adorn avenues, museums and universities across the country. "The scientific institutions they founded led to generations of disciples that continue to do the science of today," he says.

Kornblihtt carries on that tradition, in part by teaching an introductory course on molecular biology at the University of Buenos Aires. "That course has been a nursery for many young Argentine scientists," he says. It lures in many students, says Diego Golombek, a biologist at the National University of Quilmes in Buenos Aires. "Imagine that on the first day of classes, young students find themselves before the country's most well-known researcher teaching molecular biology classes with an absolutely contagious enthusiasm," he says. "He's had an influence over the new generations of biologists."

Petrillo, who has just left Argentina for a research post at the Medical University of Vienna, says that he will sorely miss the camaraderie of the tight-knit group of RNA researchers from labs and universities all over Buenos Aires. The RNArgentinos, as they call themselves, have for years organized informal seminars and get-togethers to share ideas, concerns, protocols and techniques.

Kornblihtt recognizes that Argentine scientists cannot all work in their home country and he encourages his students to "seed the world" as postdocs abroad. But he asks his university students to complete their PhDs in Argentina. "It's not necessary to leave the country to get a doctorate," he says. "We have a strong science ministry, lots of scholarships and subsidies and new research buildings. The structure to do science in Argentina is not precarious. It has many pillars." ■

VERA ROSENBERG

# HOMeward BOUND

South American efforts to repatriate scientists are paying off.

BY BARBARA FRASER

**W**hen Andrea Bragas left Argentina in 2000 for a postdoctoral fellowship at the University of Michigan in Ann Arbor, she did not know where she would eventually end up. Although the terms of her fellowship obliged her to return home, Argentina's economy was heading for a crisis and there was no guarantee of continued government funding, much less a job when she came back.

But the gamble paid off. By 2004, Argentina's economy had started to rebound and the president was pledging new investments in science and technology. Bragas returned to teach at the University of Buenos Aires and is now a nanoscientist at CONICET, Argentina's National Scientific and Technical Research Council.

Across South America, thousands of researchers have similar stories. Countries that saw some of their most promising scientists flee during decades of dictatorships or economic crises are now reversing the brain drain, luring researchers home with offers that range from short-term teaching and research fellowships to fully equipped labs and competitive salaries.

"Unlike financial capital, which is hard to recover once it has left the country, intellectual capital returns with interest," says Lino Barañao, Argentina's science and technology minister. "A scientist who has spent some years outside the country has training, networks of contacts and access to top institutions — and from a productivity standpoint can be more valuable than one who has stayed in the country."

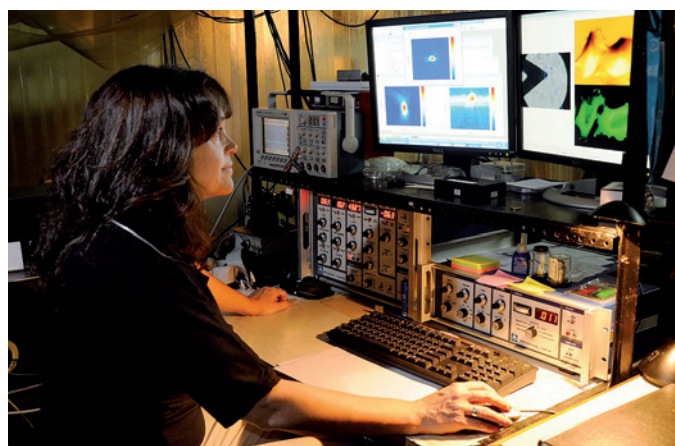
Brazil was one of the first South American countries to invest in building a base of researchers. When Lindolpho de Carvalho Dias attended the first Brazilian Mathematics Colloquium as a student in 1957, he was one of about 50 participants in a country that had few universities and no graduate programmes in science.

But the government was taking major steps to close the education gap. In the early 1950s, it created the National Council for Scientific and Technical Development (CNPq) and launched a higher-education campaign. Since then, Brazil has paid to send students abroad for graduate study, with the commitment that they would come back to teach and do research. Many of those who returned became staff members in new graduate programmes and the country has ramped up its production of scientists and engineers. The number of doctorates awarded in those fields per year nearly doubled between 2001 and 2011.

As a measure of the country's scientific growth, the mathematics colloquium currently draws about 1,000 participants a year. And research institutes in Brazil now attract both home-grown and foreign talent, adds Dias, who has served as director of the CNPq and as executive secretary of the Ministry of Science and Technology.

Like Brazil, Argentina has long sent students abroad for graduate education. But the country has only recently devoted sustained and coordinated funding to provide opportunities for returning researchers like Bragas. The science ministry now runs a programme called RAICES ('Roots') to encourage researchers to return home with offers of fully equipped laboratories and salaries comparable to those in the United States and Europe.

So far, 1,062 Argentinean scientists have returned. Most have gone to



Andrea Bragas in her nanotechnology laboratory.

public universities or research centres, although Barañao expects that to change as Argentina's private technology sector cranks up. The employer usually provides laboratory facilities, and RAICES pays moving costs and subsidizes salaries for a few years. As an added incentive, it also helps with placements for spouses.

In Chile, the Millennium Scientific Initiative — launched in 1999 — has set up centres of excellence and offers study-abroad fellowships with a commitment to return home to work. It has also established a programme called ChileGlobal, which lets Chilean scientists network at home and abroad through seminars and other activities.

Countries with smaller science budgets are also experimenting with ways to repatriate researchers through fellowships, networking and incentives. In March, Colombia's Department of Science, Technology and Innovation announced the US\$9-million 'It's Time to Return' repatriation programme. The initiative offers research posts in various fields, and hopes to lure back 500 Colombian PhD holders in its first two years.

Although brain-drain-reversal programmes take different forms, Barañao says that the key is to harness the expertise, contacts and experience of researchers outside the country — many of whom were educated at least partly at the taxpayer's expense — while expanding research facilities and opportunities at home.

Ultimately, the long-term success of these efforts may depend on the willingness of governments and companies to increase research investments, which have been climbing only modestly relative to gross domestic product in most South American countries. "You have to create a competitive research environment with top-quality, interdisciplinary research centres," says Barañao. "Even if you offer a good salary or pay relocation expenses, without those conditions, a good researcher won't return." ■

Barbara Fraser is a freelance writer in Lima.

UNIV. BUENOS AIRES EXACTAS COMUNICACIÓN



# COMMENT

**SOUTH AMERICA** Fellowship programme turns brain drain into brain gain **p.213**



**CULTURE** Half a century of art that grapples with science **p.216**

**FOOD** In conversation with Jo Robinson, phytonutrient champion **p.217**

**SYNTHETIC BIOLOGY** Luminaries petition for respect and togetherness **p.218**

ILLUSTRATIONS BY JASIEK KRZYSZTOFIAK/NATURE



## Architects of South American science

Ten research leaders call for policies to build science, and ways to build science into policy.

### ARGENTINA Strengthen networks

Eduardo Arzt is director of the BioMedicine CONICET-Partner Institute of the Max Planck Society, Argentina

Regional and cross-continental networks strengthen science in South America. They encourage young scientists to return home, motivate governments to invest in their own science, and fill gaps in core technologies such as advanced microscopy and proteomics, which require sophisticated instruments. A number of initiatives in recent years illustrate several creative approaches.

One model relies on partnerships with other prestigious institutes. For example, Uruguay's Pasteur Institute in Montevideo was founded in 2004 through an agreement with its counterpart in Paris, and the Biomedicine Research Institute of Buenos Aires, inaugurated in 2011, is a partner institute of the German Max Planck Society. Both institutes have recruited dozens of young researchers and built dedicated laboratories. They have also appointed international boards of scholars to offer advice and evaluate the quality of the science. This positive feedback loop should motivate similar evaluation schemes across other institutions.

Other programmes also foster collaborations between scientists in South America and scientists in North America and Europe. In April, Argentina became an associate member state of the European Molecular Biology Laboratory (EMBL). Symposia have already been organized, and Argentinian scientists now have access to the EMBL's state-of-the-art resources.

The Millennium Science Initiative (active in Chile and Brazil), the US National Institutes of Health's Fogarty International Center, the Howard Hughes Medical Institute, the Pew Charitable Trusts (see ►



**SOUTH AMERICAN SCIENCE**

A *Nature* special issue  
[nature.com/southamerica](http://nature.com/southamerica)

► page 213) and the Partner Groups of the Max Planck Society all sponsor individual scientists to help to create a critical mass in fields such as molecular biology, neuroscience and nanotechnology.

Regional entities have recognized the benefits of such programmes. The multinational South American trade group MERCOSUR, through its fund FOCEM, provided US\$7 million to build a biomedical research network spanning six institutions in Argentina, Uruguay, Brazil and Paraguay. The network will foster research, training and technology transfer in molecular medicine. National governments will chip in a further \$3 million.

These networks are building momentum in the region's science. As they begin to bear fruit, the time is right to build on them and not become complacent.

## PERU

### Build research capacity fast

*Gisella Orjeda is president of the National Council for Science, Technology and Technological Innovation, Peru*

It is an exciting time for science in Peru. After years of neglect, the budget of the National Council for Science, Technology and Technological Innovation (CONCYTEC) has grown 20-fold in just 18 months to almost US\$110 million, and it will continue to grow at the same rate. For the first time, Peru has a president who is prioritizing science and innovation. Journalists are trying to grasp and explain new concepts.

Now Peru needs highly qualified scientists and scientific managers. We must learn how best to organize calls for proposals, allocate funds, build programmes and reach companies. Then we must work out how to build prosperity with our new-found knowledge.

CONCYTEC establishes and promotes national policies for science, technology and innovation, and funds research. We work with local governments, the private sector, scientific institutes, universities and colleges. This is a big task for an organization of 148 people that until 2012 had an annual budget of just \$6.3 million and almost no information about the set of institutions that produce, transfer and use knowledge.

We are building these capacities: defining evidence-based policies and priorities, adhering to conflict-of-interest guidelines,

and establishing a merit-based review of proposals and incentives for innovation. We are eliminating rigid rules for immigration, buying scientific equipment and hiring qualified personnel.

I returned to Peru eight years ago, after spending ten years in France, because I wanted to make a difference in my country. After publishing the potato genome in *Nature* in 2011, I never imagined that I would have to leave science to lead science, but I have no regrets. It is thrilling to be at the helm of CONCYTEC as we face the formidable challenge of constructing a knowledge-based economy.

## BRAZIL

### Boost pro-forest economics

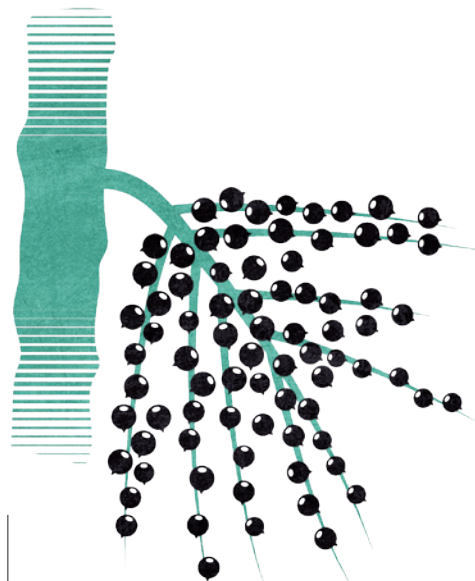
*Carlos Nobre is national secretary for research and development policies at the Ministry of Science, Technology and Innovation of Brazil*



The deforestation of the Amazon must stop: when forests are cleared for agriculture, cattle ranching and logging, the damage is felt environmentally, economically and socially. But simply curbing deforestation is not enough: sustainable-development strategies must also improve well-being for local communities.

Unfortunately, the global economy places a higher premium on meat and soya beans than on forests. Creating a new economic model for the Amazon forest will therefore take two transformations; both require science.

One strategy is to add value to locally harvested products. A good example of such a bioindustry is the açai fruit of the palm tree *Euterpe oleracea* that grows in the Amazon. Until around 20 years ago, the dark berries were a food staple consumed only by the local population. Today, açai fruit is used in produce including food, nutritional supplements, cosmetics, dyes and industrial oils around the world. Annual pulp production exceeds 200,000 tonnes and contributes more than US\$2 billion to Brazil's economy, second only to beef and tropical timber.



Local açai producers can make more than \$1,000 per hectare in annual profit, 5–10 times more than from soya and at least 15 times more than from cattle. Embrapa — the Brazilian Agricultural Research Corporation — has used açai to produce a dye for bacterial plaque that is now ready for commercial use in toothpaste and mouthwash.

More research is needed to identify uses for new and known natural products, and to scale up production. In a decade or two, it should be feasible to increase the exploitation of dozens of forest products.

A second strategy is to make better use of the large areas of already cleared forest — estimated at more than 750,000 square kilometres in the Brazilian Amazon alone — to reduce the need to clear even more. A nationwide Low Carbon Agriculture Program aims to more than double cattle occupancy per hectare within a decade. Field research conducted by Embrapa and the Brazilian cosmetics company Natura showed that oil-palm plantations on smallholdings could be integrated with other crops, such as nitrogen fixers, to obtain yields comparable to those of large-scale plantations.

Both these transformations require educating the rural and urban populations

**“In a decade or two, it should be feasible to increase the exploitation of dozens of forest products.”**

to change their ways. Technical programmes to increase agricultural productivity must reach hundreds of thousands of farmers. Isolated traditional populations will need help to reap value from collecting and selling products of biodiversity. Doing so will rely on modern communication — a new government-owned telecommunications satellite is set to start operations in 2016 to bring high-speed Internet to communities in the Amazon.



## CHILE

## Empower coastal research

Juan Carlos Castilla is professor emeritus at the Pontifical Catholic University of Chile

Rich countries can protect vast areas of their seas. Australia bans fishing in 345,000 square kilometres of the Great Barrier Reef; California protects about 16% of its coastal waters, some 2,200 km<sup>2</sup>. This approach will not work in the parts of the developing world where people's livelihoods depend on coastal fishing. A promising alternative is community-centred stewardship, boosted by research and education.

The Chilean government grants coastal communities exclusive territorial use rights in fisheries (TURFs) to extract seafood from a designated area, in exchange for a management plan that limits the annual catch proportion of algae and benthic organisms (bottom-dwelling animals including molluscs, shrimp and crabs). Around 500 of these co-management areas encompass more than 1,100 km<sup>2</sup>. The areas are only 4–10 km apart, so larvae and young animals from one area can disperse into another.

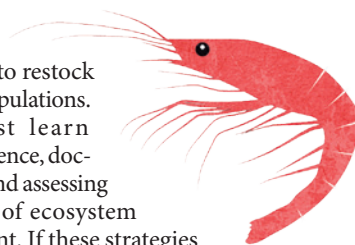
This system of fishery co-management was established in 1991. Communities differ in their performance, but results reported in 2012 revealed a desirable by-product: TURF areas show robust increases in the biodiversity of invertebrates, algae and rockfishes compared to uncontrolled areas.

Co-management empowers people to care for their resources. If a port or power plant begins operations nearby, communities demand that any damage to their area is assessed and compensated.

In unmanaged areas, the coast is overfished. TURFs are not enough. One strategy is to develop communal-management approaches for specific resources in the areas that can be fished by anyone. Regulations in Chile that came into force last year will set a total allowable catch for key species, attempting to account for a marine stock's reproductive, growth and mortality rates. A network of no-take areas between TURFs would also help. Ocean life in the no-take areas

could help to restock depleted populations.

We must learn from experience, documenting and assessing the effects of ecosystem management. If these strategies fall into place, communities can continue to fish, protect biodiversity and safeguard coastal ecosystems.



## ARGENTINA

## Fuel public-private consortia

Lino Barañao is Minister of Science, Technology and Innovative Production, Argentina

After a decade of policies aimed at boosting research, science in Argentina is starting to have positive effects on economic development and society. Now, greater involvement from the private sector is required.

Five years ago, the Argentinian government launched the Sectoral Funding Strategy to promote public-private consortia. From 2008 to 2013, more than 5,000 companies, including 80 start-ups, received a total of US\$800 million as grants or loans with below-market interest rates. The government also created programmes for postdocs and established researchers to gain experience in private companies. The number of scientists in industry increased from 7,200 in 2003 to 12,300 in 2012, and is expected to rise to more than 18,000 by 2020.

Projects funded by the strategy must combine a key enabling technology (such as biotechnology, nanotechnology, or information and communications) with a strategic area (such as health, energy, or environment and social development). They must also provide a business plan to bring an innovative product or service to market within five years. Some projects have already moved beyond proof of concept, including production of human growth hormone in the milk of transgenic cows and nanotechnology systems for drug delivery. Another example is Satellogic, a company that is developing nanosatellites for imaging. It is about to launch its third prototype and has already received private investment.

In 2012, Argentina's national research council, CONICET, and its national petroleum company, YPF, came together to create a joint company called Y-TEC. The firm, which employs more than 70 researchers, is developing technologies to exploit unconventional oil such as shale and renewable energy, and has already submitted six patent applications, three of which are licensed.

In developing countries, the science and technology sector cannot focus only on cutting-edge technologies; it must also promote social inclusion. The latter is illustrated by the Guanaco Project in the Andes, which is developing textiles for the 'responsible luxury' market. Guanacos, close cousins of llamas and vicunas, produce a fibre superior to cashmere.

In the past, science had only a cultural role in Argentina. Now it is contributing to a knowledge-based economy as a means to achieve a more just society.

## BRAZIL

## Reward quality not quantity

Sidarta Ribeiro is director of the Brain Institute at the Federal University of Rio Grande do Norte, Brazil

In the past decade, the Brazilian government has put substantial resources into education and science. It has: established a minimum wage for school teachers; allocated 1.2% of the gross domestic product to fund research; and launched the Science without Borders scholarship programme to attract foreign talent to the country and

to help promising Brazilian researchers to train abroad.

Two of the biggest remaining barriers to improving the nation's research are performance evaluation and rewards. Valuing quantity over quality is so ingrained in Brazil's scientific culture that it is nicknamed *numerologia* (numerology), a pun on the mystical belief in the power of numbers.

The official Qualis system for the evaluation of scientific papers and journals — which carries heavy weight in grant and job applications — encourages Brazilian researchers to publish as many papers as possible, regardless of the international impact of their research. Qualis does recognize different tiers of journals, but the categories are so broad as to be almost meaningless — a paper published in a journal such as *Nature* or *Science* and one in a highly specialized journal might be counted equally. Rather than gathering a full set of experiments into a coherent story, scientists gain more recognition in the system by breaking related work into multiple papers.

Independent international evaluations at universities and research institutes might be the key to rewarding innovation and cutting-edge science more effectively.

**"Independent international evaluations at universities and research institutes might be the key."**



## VENEZUELA

## Respect science and scientists

Claudio Bifano is president of the Venezuelan Academy of Physical, Mathematical and Natural Sciences

Much of Venezuela's technology and scientific capacity, built up over half a century, has been lost in the past decade. We need to restore respect and funding to basic research to halt the brain drain and reverse this catastrophic trend.

In recent years, Venezuela has invested more than 2% of its gross domestic product (GDP) in science and technology, and boasts a workforce of about 13,000 scientific researchers. But the number of publications in international journals declined by 40% in 2008–12, from roughly 1,600 to 1,000. The total number of publications in 2012 matched that of 1997, when the country had fewer than 3,500 researchers, and a science and technology budget of just 0.3% of the GDP.

According to a 2011 survey, 51% of Venezuelans over 25 years old living in the United States have finished university (compared to 13% of the US Hispanic population and 29% of all US residents). The online publication *Piel-Latinoamericana* reports that 1,100 of 1,800 physicians who graduated from medical school in Venezuela in 2013 have left the country. In other words, educated Venezuelans are fleeing — or are being forced out. For example, in 2003, roughly 1,000 professionals, mostly physical scientists and engineers, were fired from Venezuela's petroleum research and development institute, INTEVEP. International agencies report that no patents have been granted since that time.

Since 1999, the Venezuelan government has imposed a political model called socialism of the twenty-first century. I and others find it based mainly on authoritarianism, with some ideas from Marxist philosophy and extreme populism. Science, according to the minister for science and technology, is for the solution of societal problems. The National Science, Technology and Innovation Plan (2005–30) says that science must be conceived as a process that involves new participants, such as the holders of traditional and local knowledge.

To achieve this goal, the Ministry of Science, Technology and Innovation supports projects submitted not only by scientists but also by those without scientific training and by organizations such as community councils, environmental groups and associations geared towards the social services. Funded programmes include one that

distributes computers to school children and missions for a remote-sensing satellite and a data-transmission satellite. These may be laudable projects, but they are not science.

Allowing those who lack scientific training to access public funds for scientific research trivializes science.



## BRAZIL

## Banish bureaucracy

Jose Eduardo Krieger is provost of research at the University of São Paulo, Brazil

Brazil needs a better environment for knowledge creation and innovation. Bureaucracy currently holds back research. Fixing this will require changes to institutional policy and national legislation.

At the University of São Paulo, for instance, we began a major initiative in 2011 to enable scientists to focus on what they do best, rather than wasting time filling in forms. The university is the largest research institution in South America, responsible for about 20% of all papers published in Brazil every year. The institution's 6,000 scientists win almost half of the US\$450 million that the state of São Paulo awards to support research.

But most Brazilian grants do not cover overhead or indirect costs, such as facility maintenance. So our universities lack the support offices that North American and European researchers rely on to help with ordering equipment and reagents, paying invoices, financial reporting, contract negotiation and account monitoring. Every researcher must set up these systems individually.

By the end of this year, the University of São Paulo will roll out a digital platform to assist researchers with procurement, accountability and operations. We are also creating a network of trained project managers to assist specialized schools and large research groups. These measures follow a \$100-million, four-year effort by the university to reorganize its research enterprise.

More than 100 research support groups have been created, each with a technician, to encourage scientists to organize themselves into interdisciplinary clusters.

These strategic moves will be complemented by improvements in the regulatory laws currently under discussion in the Brazilian Congress. These should allow equipment and consumables for academic research to be imported more quickly and easily — giving our scientists more time for research, and helping them to compete with their peers in North America and Europe.

## CHILE

## Base policy on evidence

Pablo C. Guerrero is assistant professor at the University of Concepción, Chile; Mary T. K. Arroyo is director of the Millennium Institute of Ecology and Biodiversity, Chile

Chile needs a system for formulating public policy on the basis of sound scientific information. The government's decision in March not to create a ministry of science passed up a valuable opportunity for that.

The current disconnect between science and policy within the government is worrying, as two recent examples show. First is the devastating fire that swept through

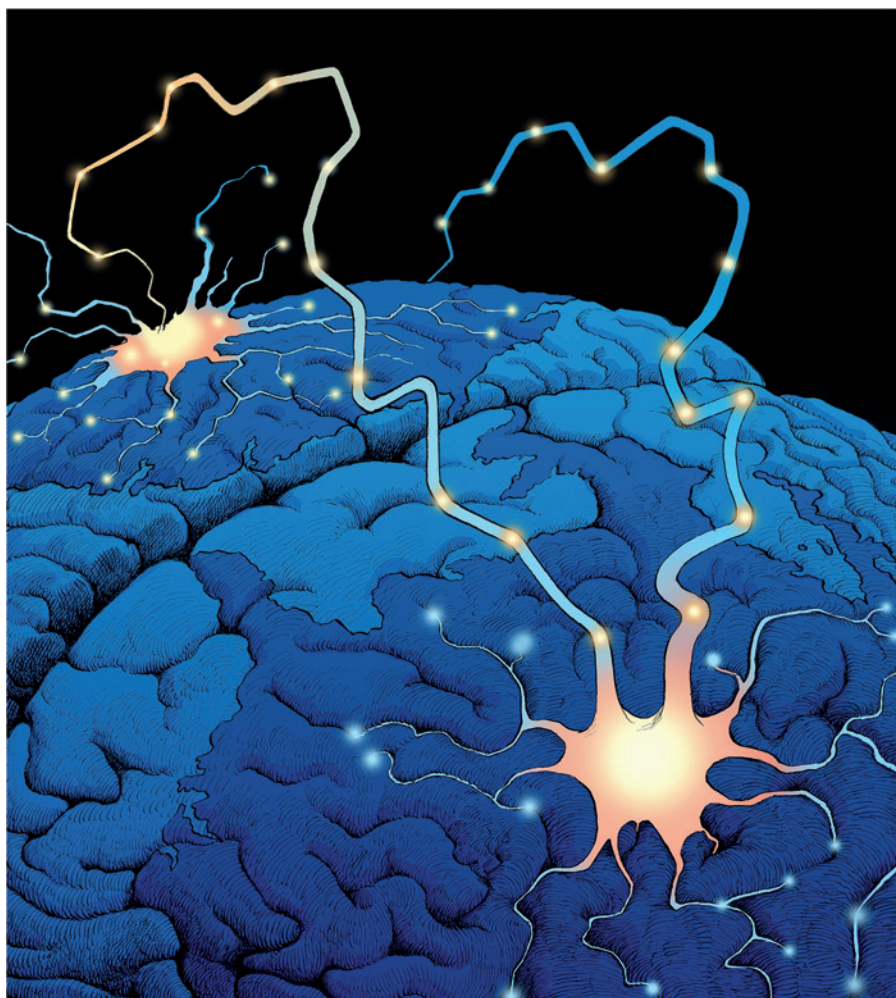
**"The current disconnect between science and policy within the government is worrying."**

parts of the city of Valparaíso in April. For decades, authorities ignored ecologists' warnings about expanding highly flammable eucalyptus plantations that are now near many cities in central Chile, and where the Valparaíso fire started.

Second, Chile has experienced seven earthquakes measuring magnitude 7 or more in the past decade. Here, too, scant attention was paid to scientists' predictions about the accumulation of seismic strain.

Some have suggested that Chile's highly regarded science-funding body, the National Commission for Scientific and Technological Research (CONICYT), could regain its past influence and advise on public policy once more. To do so it will need to adjust its current emphasis on impact factors and international recognition of basic science. CONICYT should give explicit credit to basic-science problems that are relevant to the concerns of Chileans, such as the availability of water resources in a changing climate and innovative ways to use minerals. ■





# Turning brain drain into brain circulation

Overseas scholarships that encourage scientists to return to their home countries are helping to rebuild science in Latin America, says **Torsten Wiesel**.

**I**t takes a long time for a country to build a strong base in science, but only a short time to destroy it. Germany was a sad example. It was a world leader in the sciences for more than a century, until its science base was demolished during the Nazi era, and the country ceded its position to the United States. It has taken decades for Germany to rise again to its current level of excellence.

The German experience has much in common with the situation in Latin America, where authoritarian regimes came to power in the mid-twentieth century in countries including Brazil, Chile and Argentina. As a consequence, many of the continent's best scientists emigrated to the United States, Europe and Canada. When the dictatorships were finally shaken off in the 1980s and 1990s, the departed scientists were settled

in their new homes and had little incentive to return to countries left laden with debt.

Many have forgotten that science in Latin America was once robust. For example, Bernardo Houssay, who won the 1947 Nobel Prize in Physiology or Medicine, directed the Institute of Physiology at Buenos Aires University until 1943, when the government fired him for advocating for democracy; his protégé, Luis Leloir, won the 1970 Nobel Prize in Chemistry. Several emigrants also became laureates, including the immunologist Baruj Benacerraf, from Venezuela, and the biochemist César Milstein, from Argentina.

Against this background, the Pew Latin American Fellows Program was founded to help to rebuild and strengthen biomedical sciences in the region. From its inception, the programme has been linked to the pre-existing Pew Biomedical Scholars Program, which each year provides around 20 promising newly independent US scientists with four-year scholarships, funded by the Pew Charitable Trusts, a non-profit organization based in Philadelphia, Pennsylvania.

In March 1989, at the annual meeting of the scholars programme in Puerto Vallarta, Mexico, a group of these scholars — struck by the lack of resources of their counterparts in Mexico — sought help from Rebecca Rimel, president of the Pew Charitable Trusts. Later, Rebecca and I discussed the best ways to train talented students from Latin America, and our ideas crystallized into the fellows programme.

## REPATRIATION RATES

Since the founding of the Pew Latin American Fellows Program in 1991, about ten graduate students each year have been awarded two-year postdoctoral fellowships to work in some of the best labs in North America. It is no surprise that some remain abroad to continue their careers in more developed countries. What is surprising is that more than 70% return to their home countries, which may not always allocate sufficient resources to cutting-edge research (see 'Bringing science home'). For comparison, the Human Frontier Science Program, a multinational initiative that supports the life sciences, also funds postdoctoral fellows worldwide — but fewer than half of those who train in the United States return to their home countries.

Pew fellows who remain in North America have positions in leading universities and several have established joint projects with labs in their home countries, as well as hosting new fellows. The annual



## PEW LATIN AMERICAN FELLOWSHIP

*Bringing science home*

Becoming a great scientist requires exposure to greatness.

At a 1997 orientation meeting in Costa Rica for new postdocs, Torsten Wiesel, the co-founder of the Pew Latin American Fellows Program, told us that the best scientists are not necessarily more creative or smarter than everyone else, but that they had the opportunity in their junior years to conduct and discuss science in prime environments.

I earned my PhD in 1996 from the University of Chile in Santiago, studying how ions move through proteins extracted from neurons. I wanted to apply that work in living brains. Senior members in my department told me about the Latin American fellows programme and helped me to find a postdoctoral adviser.

Charles Zuker, then at the University of California, San Diego, accepted me into his lab and taught me to study how flies sense the world. It was an amazing experience to be in the Zuker lab when seminal work on taste and pressure receptors was happening. I was part of the team that helped to show how the organization of proteins in photoreceptor cells is essential for flies to see light. I returned home to work as a junior professor at the University of Chile in 1998.

Even now, few institutions in South America provide start-up funds to new faculty members. Most young professors have to join senior laboratories or sit in an empty lab, sometimes for more than a year, before getting their first grant. By contrast, I had a US\$35,000 repatriation fund from my Pew fellowship. The money was enough

to buy small, essential equipment to start doing some simple experiments soon after I returned: a table-top centrifuge to separate cells into basic components, power supplies, electrophoresis chambers to run gels for DNA analysis, a mechanical shaker to grow bacteria and some reagents.

Since then, I have trained nearly two dozen students to work with flies and have helped four researchers to set up their own labs for fly research in Chile. I have also directed three international courses to train Latin American students to use the insects (and, more recently, worms) as animal models.

And my relationship with Pew continues. I have started collaborations with scientists from other countries whom I met at annual Pew alumni meetings. For the past five years, I have served on the regional Pew committee that selects six Chilean candidates for the fellowship. We look for young researchers who have connected with a great lab and proposed adventurous projects — particularly to work in areas or with animal models that are not available at home. The hope is that they will bring those skills back to their native countries.

Chile has an 80% repatriation rate. That bespeaks both a good selection process and the importance of the start-up money for returning fellows. Scientific agencies and governments in Latin America should try to replicate these measures to help to build a stronger and more innovative scientific community. **Jimena Sierralta,** [University of Chile](#)

meetings are attended by Latin American fellows, biomedical scholars and senior advisers, including Nobel laureates and Howard Hughes Medical Institute scholars. Participants share ideas and start collaborations as a result of the meetings.

**SUCCESSFUL SCHOLARS**

In a survey sent out in 2013 to 202 alumni of the Latin American fellows programme between 1991 and 2011, an impressive 151 responded. Alumni who have returned to their home countries include department heads and university provosts. Nearly half reported holding a director position, such as department chair or head of an academic discipline. On average, each fellow had published 15 papers, and those who had returned home had trained 13 scientists, from technicians and graduate students to visiting scholars.

Last month, the journal *Cell* highlighted a 2003 Pew fellow, immunologist Dario Zamboni, as one of 40 notable scientists under 40 years old. Zamboni is head of the Innate Immunity and Microbial Pathogenesis laboratory at the University of São Paulo in Brazil. His group is working out how the body responds to intracellular parasites, including the one that causes Chagas disease — a problem in poor, rural areas of South America. Doing science in Brazil involves hurdles that would not exist in the United States, but he is determined

to improve the system for other scientists in the country.

Selection of fellows starts with established researchers in Latin America. Argentina, Brazil, Chile and Mexico have national committees of former Pew fellows and senior scientists. Each committee selects six applicants by evaluating research proposals and interviewing a dozen or so of the most promising students. (The chairs of these committees act together as a fifth multinational committee for applicants from the other countries in the region.)

***“The fellows programme is just a drop in the ocean relative to the need of the entire continent.”***

Thirty applications are chosen in total to be evaluated by a central committee of outstanding US scientists with strong ties to Latin America. Several are emigrants from the dark periods in their countries of origin. These committees work hard to select the most promising scholars and send them to the best labs.

The Brazilian state of São Paulo plans to augment the benefits that are open to returning Pew fellows: they can apply for a generous four-year stipend to get their new labs off the ground. The hope is that other nations will use their own resources to extend this initiative to foster their best scientists.

The absolute number of Latin American fellows is small — fewer than 250 in a region with more than 400 million people. But my impression is that they have an outsized influence, shaping expectations of what it means to be a scientist in Latin America, and the fellows' high expectations of themselves.

That said, the fellows programme is just a drop in the ocean relative to the need of the entire continent. This is perhaps especially true now that larger programmes exist in several Latin countries to support the training of scientists abroad and to encourage trained scientists to return home, such as the Brazil Scientific Mobility Program (see page 207).

Nonetheless, like a seed planted in a fertile soil, the Pew programme has flourished over the past 20 years. The plant will no doubt continue to grow and to support its ecosystem. The ultimate success would be that this type of programme is no longer needed because each country would have developed strong, independent scientific establishments. But for now, we need to bolster the support for scientists in emerging countries, in Latin America and elsewhere. ■

**Torsten Wiesel** is president emeritus of Rockefeller University in New York City, USA. He won the 1981 Nobel Prize in Physiology or Medicine.  
e-mail: [wiesel@rockefeller.edu](mailto:wiesel@rockefeller.edu)





Robert Rauschenberg's 1963 installation *Oracle*, created with engineers Billy Klüver and Harold Hodges.

## ARTS

# The third culture

**Michael John Gorman** is intrigued by a survey of art informed and invigorated by science.

After months of injections with horse immunoglobulin in 2011, artist Marion Laval-Jeantet had a transfusion of horse blood in a Ljubljana art gallery. She walked around the donor animal on prosthetic hooves; then samples of her hybrid blood were freeze-dried and placed in engraved aluminium cases. In 2005, a New York gallery showed a starburst of glass orbs and aluminium rods depicting the explosion of space after the Big Bang, by sculptor Josiah McElheny and cosmologist David Weinberg.

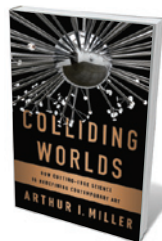
Such are the collaborations chronicled by historian Arthur I. Miller in *Colliding Worlds*. Miller argues that we are seeing the emergence of a “third culture” — a term coined by writer John Brockman — in which boundaries between art and science dissolve.

The past decade has seen a proliferation of galleries, labs and residency programmes devoted to mingling art and science. Miller surveys these, from London's Wellcome Collection to the Ars Electronica Futurelab in Linz, Austria; the Science Gallery at Trinity College Dublin (of which I was founding director); Le Laboratoire in Paris; and the Collide@CERN artist-residency programme at Europe's particle-physics lab near Geneva, Switzerland. He provides engaging pen portraits of many of the artists involved, such as

Evelina Domnitch and Dmitry Gelfand, who experiment with sonoluminescence.

Miller touches on early examples of cross-pollination, such as physicist Niels Bohr's interest in Cubism, but locates the origins of the modern art-science movement in 1966, with *9 Evenings: Theater and Engineering*. These experimental “happenings” had proto-pop-artist Robert Rauschenberg and avant-garde composer John Cage as key participants, in the presence of Andy Warhol, Marcel Duchamp and other New York art luminaries. Unfortunately, technical disasters and delays led to negative press coverage.

The unlikely hero of Miller's story is Billy Klüver of Bell Labs in Murray Hill, New Jersey, the instigator of *9 Evenings* and a gifted engineer. Klüver's earlier collaboration with Swiss artist Jean Tinguely — on *Homage to New York*, a self-destructive kinetic artwork made from bicycle and pram parts — led to meetings with Rauschenberg and



**Colliding Worlds: How Cutting-Edge Science is Redefining Contemporary Art**  
ARTHUR I. MILLER  
W. W. Norton: 2014.

other high-profile artists. Klüver persuaded his Bell Labs colleagues that artists would stimulate new directions in technology. As the philosopher Marshall McLuhan put it in his 1964 book *Understanding Media*, art can be considered “precise advance knowledge of how to cope with the psychic and social consequences of the next technology”.

The twenty-first-century explosion of art-and-science programmes and spaces has been fostered partly by significant investment from funding organizations. Miller documents the tension here between art's roles in illustrating, communicating and interrogating science.

Some artists, including Antony Gormley or McElheny, draw on areas such as foam physics or cosmology as aesthetic inspiration. Others provoke critical discussion around future directions of research, such as Australian performance artist Stelarc, who had an ear-shaped scaffold implanted into his forearm. Discussing the rationale of the Collide@CERN residency, which has featured sound sculptor Bill Fontana and photographer and media artist Julius von Bismarck, CERN director Rolf-Dieter Heuer tells Miller that revealing what scientists are doing for society is key. “To transmit that through art ... opens horizons,” says Heuer. It is a suggestion that would sit uneasily with many of the critically engaged artists whom Miller discusses.

Experimental art-science collaborations have not always been embraced by conventional galleries and collectors. Peter Weibel, founding director of the Centre for Art and Media (ZKM) in Karlsruhe, Germany, tells Miller that this should not be a concern because “private industry will finance” artists, liberating them from the vagaries of the market. Indeed, *Colliding Worlds* opens with the heady atmosphere of Bell Labs half a century ago; towards the end, it considers digital artists Scott Draves and Aaron Koblin, who both work for Google.

Unlike other surveys, such as artist Stephen Wilson's *Art + Science Now* (Thames and Hudson, 2010), *Colliding Worlds* features interviews with the artists, scientists and engineers involved in projects from speculative design to data visualizations, sound art and cosmetic surgery. Such tales enliven the book. But it is hard to accept “artsci”, as Miller terms it, as a coherent movement. The third culture, he shows, consists instead of exciting, experimental and mutually enriching collisions.

Ultimately, Miller suggests, such collisions — once in the mainstream — become just ‘art’. The important question that remains is whether such art can alter the direction of scientific research, beyond provoking public discussion and debate. ■

**Michael John Gorman** is chief executive of Science Gallery International in Dublin.  
e-mail: michaeljohn.gorman@sciencegallery.com

MUS. NATL D'ART MODERNE, CENTRE POMPIDOU, PARIS/PETER WILLI/BRIDGEMAN IMAGES/© ESTATE OF ROBERT RAUSCHENBERG, DACS, LONDON/VAGA, NEW YORK



Jo Robinson with vegetables from her demonstration garden in Washington state.

## Q&A Jo Robinson

# The nutrient hunter

*Investigative food journalist Jo Robinson has spent more than a decade scouring the literature on plant nutrition. Her demonstration garden in Washington state opens this month as her book *Eating on the Wild Side* (Little, Brown, 2013) emerges in paperback. She talks about eating tomatoes to protect from sunburn, why bitter is better — and how purple is the new green.*

### What is the thesis of *Eating on the Wild Side*?

All plants make phytochemicals, which protect against predation, disease and other threats. When we consume certain plants, we may also receive some protection — I give evidence from the scientific literature in my book. Lycopene in tomatoes guards against ultraviolet light, and has been shown to protect us from sunburn, for example. A relatively new discovery is that since the invention of farming, we have been breeding varieties with progressively fewer beneficial phytochemicals, partly because many taste bitter or astringent. Fruit and vegetables with fewer of these compounds may offer less protection against hypertension, cholesterol, inflammation and other ills. Part of my work is to identify heirloom and modern varieties, such as Purple Peruvian potatoes, that are rich in phytochemicals and pleasing to the palate.

### Which fruit and vegetables should we eat?

Purple, blue, red or black plants such as most berry varieties and red cabbage are good choices because they contain a

family of pigments known as anthocyanins. Test-tube, animal and now a few small-scale human studies show that anthocyanins have the potential to curb the risk of cardiovascular disease by reducing inflammation, improving blood lipids and lowering blood pressure. A pilot study determined that anthocyanin-rich berries slowed the growth of cancer cells in people with colon cancer. Evidence is mounting that anthocyanins may also slow the decline of cognition and memory that accompanies old age.

### Is colour the only indicator of such effects?

No: most phytochemicals are not highly pigmented. The drab globe artichoke has more antioxidant activity than more brightly coloured vegetables because of its high concentration of colourless cynarin (which increases bile secretion and may protect the liver from carcinogens) and chlorogenic acid, which has antihypertensive effects. White onions, leeks and shallots contain an anticancer and flu-fighting compound called quercetin. Some varieties of white-fleshed

peaches have more antioxidant activity than yellow-fleshed peaches, even though the yellow varieties have more of the pigmented phytochemical  $\beta$ -carotene.

### What is the best way to preserve the nutrients in stored vegetables?

Once a plant is harvested it does not die immediately. The harvested part is still metabolically active and begins to burn its natural sugars and lose phytochemicals and flavour. You can slow this process by reducing its exposure to oxygen, storing it in the fridge in a sealed plastic bag with 10–20 pinpricks.

### What are the best ways to cook vegetables?

I cringe when I see people boiling vegetables, because the cells burst and nutrients leach out into the water. Lightly sautéing in oil is fine, but steaming is almost always best because it reduces exposure to water. If you microwave an ear of corn in the husk, you preserve nutrients and taste. Microwaving is also best for thawing berries, because it destroys an enzyme called polyphenol oxidase that breaks down antioxidants.

### Could biotechnology help us to breed more-nutritious plants?

In my view we will never achieve the nutrient content of phytochemical-rich foods through genetic engineering. Say that we find a gene that produces cabbage with more cancer-fighting glucosinolates. This family of health-enhancing compounds is only one out of dozens in the vegetable, and, ultimately, it may not prove to be the most beneficial. But there is great promise in crossing wild species with modern ones through conventional breeding, which introduces myriad genes. In my garden I grow hybrid blackberries called Wild Treasure that are thorn-free and highly productive, but retain the nutrition and luscious flavour of wild berries.

### Why open a demonstration garden?

I want to show people that by growing their own food they can choose varieties that will increase their odds of living longer and healthier lives. In my own garden there is a wild crab apple from Nepal that has more antioxidants in a single teaspoon than a large Honeycrisp apple. There are Indigo Rose tomatoes, an inky black variety rich in anthocyanins. And there are purple varieties of carrots, cauliflower and asparagus. You could say that purple is the new green.

### What is next for you?

I may write a cookbook about findings on how to preserve and enhance plant phytochemical content. I also have enough data to write a book about nutrient-rich beverages such as tea, wine, coffee, whisky and beer. ■

INTERVIEW BY JASCHA HOFFMAN



# Correspondence

## An early champion of women's rights

In his 1859 book *On the Origin of Species*, Charles Darwin argues that “all animals and plants have descended from some one prototype”. In none of the book's six editions does he refer to this common ancestor as being an animal-like hermaphrodite with male and female gonads, as Kimberly Hamlin suggests in her book on Darwinian feminism, *From Eve to Evolution* (reviewed in *Nature* **509**, 424; 2014). Hamlin writes, for example, that “the possibility of a hermaphroditic past ... opened up a new world of gendered possibilities”.

It was the co-discoverer of natural selection, Alfred Russel Wallace, who was a public advocate of women's rights. As reported in *The Times* on 11 February 1909, he wrote: “All the human inhabitants of any one country should have equal rights and liberties before the law; women are human beings; therefore they should have votes as well as men.”

**U. Kutschera** *Institute of Biology, University of Kassel, Germany.*  
kut@uni-kassel.de

## Synthetic biology: missing the point

Volker ter Meulen warns that if environmental groups and others exaggerate the risks of synthetic biology it could promote over-regulation, which he says happened for genetically modified organisms (*Nature* **509**, 135; 2014). But the point of supporting synthetic biology is not about making sure that science can go wherever it wants: it is about making the type of society people want to live in.

In the United States, for example, the rapid and uncritical introduction of genetically modified organisms prevented debate on issues such as alternative innovation pathways, and the impact on biodiversity

and pest resistance. Many believe that these issues would have been better addressed through earlier and broader public discussion of the uncertainties surrounding transgenic organisms (see S. Jasanoff *Designs on Nature* Princeton Univ. Press; 2005).

In our view, ter Meulen trivializes the role of social scientists in suggesting that they could help the synthetic-biology debate by finding better ways to communicate what scientists think. He also implies that public concern over such technologies and their governance reflects only a failure to understand the science of risk assessment — but this ‘deficit model’ of public concerns has long been discredited (see A. Irwin and B. Wynne *Misunderstanding Science?* Cambridge Univ. Press; 1996).

It is not unknown for scientists themselves to foster exaggeration and uncritical acceptance of claims, or to focus on anticipated benefits rather than on risks. This practice may be at the heart of wider public concerns about responsible innovation (see, for instance, [go.nature.com/zehdpd](http://go.nature.com/zehdpd)).

**Sam Weiss Evans** *University of California, Berkeley, USA.*  
[samuel.evans@berkeley.edu](mailto:samuel.evans@berkeley.edu)  
\*On behalf of 21 correspondents (see [go.nature.com/romzbu](http://go.nature.com/romzbu) for full list).

## Synthetic biology: a global approach

Despite some success in advancing best practices for synthetic biology in ethics, safety, security and the environment, the conversation about a global “tribal gathering” is only just beginning (*Nature* **509**, 133; 2014).

In 2006, when the field was starting to appreciate the concepts and conditions for success, the Synthetic Biology Engineering Center (Synberc; [www.synberc.org](http://www.synberc.org)) was founded with support from the US National Science Foundation. It consisted of 20 investigators who

helped to lay the foundations for synthetic biology at a time when tribalism probably still did us some good.

Synberc is now a singularly diverse tribe. It aims to examine the broad social context of biotechnology research practice through programmes that involve political scientists, legal scholars, ethicists, theologians, industrial practitioners, anthropologists and others, along with its own scientific advisory board.

We call for global expansion of the Synberc model into a more inclusive organization that is committed to advancing responsible scientific and social progress in synthetic biology. The main US funding agencies and their counterparts worldwide need to join with academics, industrial partners and society to support this long-term, internationally coordinated effort (see also V. ter Meulen *Nature* **509**, 135; 2014).

**Jay D. Keasling** *University of California, Berkeley, USA.*  
[keasling@berkeley.edu](mailto:keasling@berkeley.edu)

\*On behalf of 16 correspondents (see [go.nature.com/bp83hq](http://go.nature.com/bp83hq) for full list).

## Successes for UK cancer partnership

Your examples of important hybrid academic-industrial partnerships for drug development come mostly from the United States (*Nature* **509**, 146; 2014). The Institute of Cancer Research in London has long benefited from such hybrid models. When normalized for each faculty member, our income from intellectual property is highest among UK universities and ranked in the top ten relative to US institutions (see [go.nature.com/ohyuqj](http://go.nature.com/ohyuqj)).

Since 2005 we have discovered 17 drug candidates — in several cases with our industry partners — and 7 of these have progressed to phase I/II clinical trials. Our drug abiraterone was approved

in the United States and Europe in 2011, and has changed clinical practice for treating advanced prostate cancer (see J. S. de Bono *et al. N. Engl. J. Med.* **364**, 1995–2005; 2011). Other examples include drugs that target breast, lung and other cancers by inhibiting proteins such as phosphatidylinositol-3-OH kinase (F. I. Raynaud *et al. Mol. Cancer Ther.* **8**, 1725–1738; 2009) and the molecular chaperone HSP90 (S. A. Eccles *et al. Cancer Res.* **68**, 2850–2860; 2008).

These successes are a result of taking early academic risks, combining academic and pharmaceutical expertise, and implementing strong leadership and project management. Other contributing factors include running multiple projects on a competitive scale, establishing long-term financial support and — most important — selecting productive and timely industrial collaborations.

**Paul Workman** *The Institute of Cancer Research, London, UK.*  
[paul.workman@icr.ac.uk](mailto:paul.workman@icr.ac.uk)

## Forgotten founder of bibliometrics

Besides being one of the conceptual inventors of the Internet (P. Ball *Nature* **509**, 425; 2014), the Belgian librarian Paul Otlet first coined the term ‘bibliometrics’. In his book *Traité de Documentation* (1934), he called for the foundation of a new field, *bibliométrie*, which he defined as the measurement of all aspects related to the publication and reading of books and documents.

As an example, Otlet suggested recording how often a particular book or author is read. He noted that mathematics was becoming increasingly important in most scientific fields, including in biology and sociology, and felt that it should be included in library science as well.

**Ronald Rousseau** *KU Leuven, Belgium.*  
[ronald.rousseau@kuleuven.be](mailto:ronald.rousseau@kuleuven.be)

## Diamond gets harder

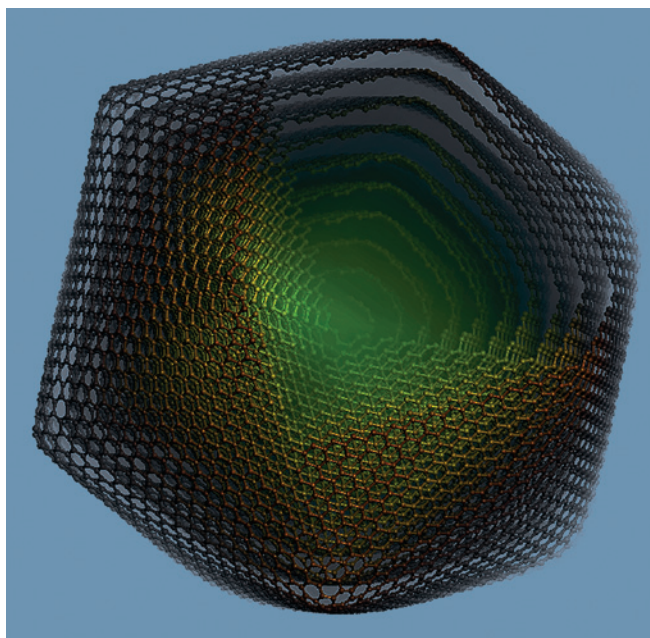
Composite materials that incorporate diamond are among the hardest in the world, but fail under extreme conditions. A nanostructured form of diamond, made from onion-like carbon precursors, might overcome this problem. [SEE LETTER P.250](#)

JAMES BOLAND

**D**iamond is a famously strong material with outstanding properties, such as high wear resistance and hardness. For this reason, it has long been used in cutting and drilling tools, but poor thermal stability has limited its application. On page 250 of this issue<sup>1</sup>, Huang *et al.* report the synthesis of 'nanotwinned' diamond, in which nanometre-scale crystals share some lattice points. The authors find that the resulting material is much harder and more thermally stable than naturally occurring diamond.

The ancient Egyptians may have been the first to use diamonds in tooling, although the evidence for this is unsubstantiated. But rock drilling with diamonds has been more reliably dated to the eighteenth century<sup>2</sup>. The need for high-strength, hard-wearing drill bits for industrial drilling and oil exploration led to the development of a new class of superhard material in the 1980s consisting of diamond grains bonded with metallic cobalt. The main disadvantage of these materials is that the cobalt catalyses the breakdown of diamond to graphite at temperatures above 700 °C. A diamond composite was developed around the same time<sup>3</sup>, in which the cobalt binder was replaced by a ceramic material, silicon carbide, and was shown to be stable under harsh and severely abrasive rock-cutting conditions to temperatures in excess of 1,200 °C. However, this thermally stable diamond composite material has yet to be widely adopted as a cutting element in tools for the mining, drilling and manufacturing industries for reasons of cost.

A major drawback of diamond-based composites has been their low fracture toughness (a measure of resistance to crack propagation), which can cause them to fail catastrophically. Harder diamond composites, which have higher concentrations of diamond, have lower fracture toughness. Nevertheless, these materials have high wear resistance, and have formed



**Figure 1 | Computer model of an onion carbon nanoparticle.** Huang *et al.*<sup>1</sup> used such nanoparticles to make an ultrahard, nanostructured form of diamond. (Image taken from ref. 1.)

the basis of long-lasting tools for industrial use, provided that the mechanical loading on them is controlled.

Hardness is not governed by composition alone; the grain size of the constituent phases of the materials is also a factor. For hard and brittle materials such as diamond composites, hardness and strength increase with decreasing grain size, as expressed by the Hall–Petch relationship<sup>4,5</sup>. Normally, such improved hardness is accompanied by a decrease in fracture toughness; this inverse relationship was a generally accepted model until nanostructured materials were thoroughly investigated for their mechanical properties. In such materials, the inverse relationship no longer holds when the grain size is less than about 100 nanometres, and fracture toughness can actually increase with decreasing grain size<sup>6</sup>. These materials, including diamond composites with constituents that have nanoscale grains, have been shown to have outstanding fracture toughness.

Grain-size-reduction techniques for improving the fracture toughness of ultra-hard

materials have proved useful, but seem to have been limited by either the material involved or the technology used. Further improvements to such materials therefore seemed unlikely, unless alloyed nanograined materials with even higher intrinsic hardness could be discovered. However, Huang and colleagues<sup>1</sup> demonstrate that a further reduction in the crucial hardness-related length scale of grains is achievable.

Researchers from the same group had previously reported<sup>7</sup> a process for making a nanotwinned form of boron nitride — a material with a diamond-like atomic arrangement. They therefore decided to mimic that process with diamond, by subjecting carbon nanoparticles consisting of concentric graphite-like shells (known as onion carbon nanoparticles; Fig. 1) to pressures in the range of 18–25 gigapascals at temperatures of 1,850–2,000 °C. The resulting transparent material consisted of nanotwinned, nanocrystalline diamond.

The hardness of Huang and co-workers' material reached about 200 GPa; for comparison, hardness values for single-crystal diamonds range from 60 to 130 GPa, and those of nanocrystalline diamonds without nanotwins are 130–145 GPa (ref. 8). Another outstanding property is its high fracture toughness, which is greater than that of other commercially available diamond composite materials. Remarkably, the nanotwinned diamond was stable against oxidation in air at temperatures above 1,000 °C — higher than the authors expected.

Huang *et al.* prepared millimetre-sized pieces of their material on a laboratory scale, but it remains to be seen whether their process can be used on an industrial scale. Success will depend in part on whether starting materials of sufficiently high quality can be made. Nanocrystalline diamond has previously been sintered — fused at high temperature and/or pressure — to manufacture anvils that are used for high-pressure, high-temperature phase studies of geological materials<sup>8</sup>, and



## PLANT BIOLOGY

# Pass the ammunition

**Tomato plants that have been damaged by herbivorous insects emit airborne chemicals that warn neighbours of an impending attack. It emerges that the receiving plants transform these signals into defensive weapons.**

MARK C. MESCHER &  
CONSUELO M. DE MORAES

similar scientific applications could be predicted for nanotwinned diamond. However, the material's creep (the tendency of a material to deform permanently in response to long-term mechanical stresses) and fatigue properties need to be measured. If the deformation mechanism changes from one that is based on crystallographic defects to one based on sliding at grain boundaries, as commonly occurs during 'superplastic' deformation when solid materials are heated, then methods for pinning grain boundaries would be required<sup>9</sup>.

Nanodiamonds have progressed over the past decade or so from being speculative curiosities to fully functioning materials useful for a broad range of applications. Individual nanoparticles consisting of only a few hundred carbon atoms arranged into the diamond structure are being used in such diverse areas as drug delivery, bioimaging and tissue generation<sup>10</sup>. Nanodiamonds, either aggregated or disaggregated in lubrication fluids, can also form low-friction interfaces that reduce wear on moving components at both the macro- and microscale<sup>11</sup>.

Equally important is the innovative and rapidly developing research on the consolidation and sintering of nanodiamonds to make solid composite materials that have a wide range of remarkable properties, such as high thermal conductivity, optical transparency, chemical inertness and high tolerance to radiation damage. These composites were initially produced on scales barely higher than that of the nanoparticles themselves, but extraordinary progress in high-pressure, high-temperature technology<sup>8</sup> now means that the materials can be produced at sizes that have applications across several industries. The incorporation of nanotwinned, nanocrystalline diamonds into composites might lead to materials that have even more extraordinary properties. ■

**James Boland** is in the Division of Earth Science and Resource Engineering, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Pullenvale, Queensland 4069, Australia.  
e-mail: jim.boland@csiro.au

- Huang, Q. *et al.* *Nature* **510**, 250–253 (2014).
- Tolansky, S. in *Science and Technology of Industrial Diamonds* Vol. 2 (ed. Burs, J.) 341–349 (Ind. Diamond Inform. Bur., 1967).
- Wilks, E. & Wilks, J. *Properties and Applications of Diamond* (Butterworth-Heinemann, 1991).
- Petch, N. J. *J. Iron Steel Inst.* **174**, 25–28 (1953).
- Hall, E. O. *Proc. Phys. Soc. Lond. B* **64**, 747–753 (1951).
- Zhao, Y. *et al.* *Appl. Phys. Lett.* **84**, 1356–1358 (2004).
- Tian, Y. *et al.* *Nature* **493**, 385–388 (2013).
- Irifune, T. & Sumiya, H. in *Comprehensive Hard Materials* Vol. 3 (eds Mari, D., Llanes, L. & Nebel, C. E.) 173–191 (Elsevier, 2014).
- Suryanarayana, C. & Al-Aqeeli, N. *Prog. Mater. Sci.* **58**, 383–502 (2013).
- Mochalin, V. N., Shenderova, O., Ho, D. & Gogotsi, Y. *Nature Nanotechnol.* **7**, 11–23 (2012).
- Ivanov, M. *et al.* *Nanosyst. Phys. Chem. Math.* **5**, 160–166 (2014).

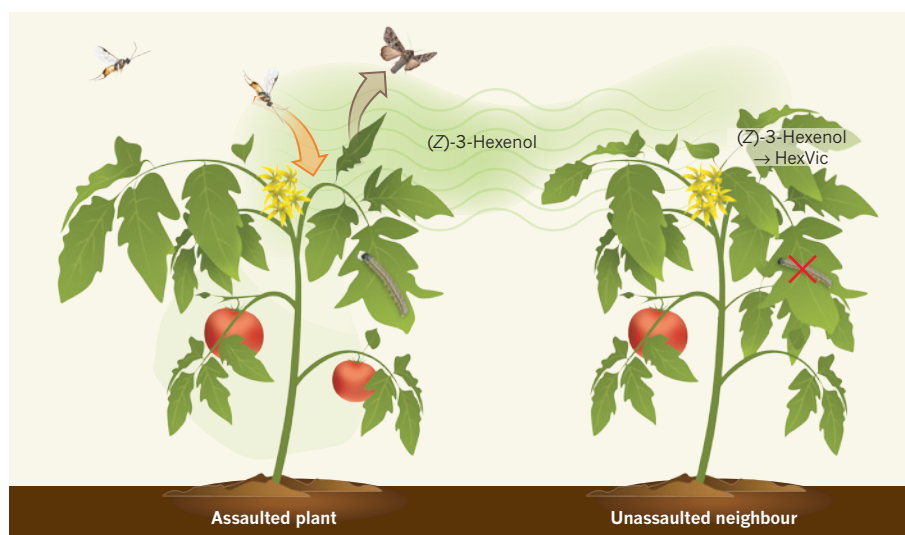
Plants may seem passive, but in fact they respond in complex ways to diverse features of their environment. It is becoming increasingly clear, for example, that plants perceive and respond to environmental odours. However, almost nothing is known about the mechanisms by which plant olfaction occurs. Writing in *Proceedings of the National Academy of Sciences*, Sugimoto *et al.*<sup>1</sup> report that when plants are exposed to odours emitted by neighbours that have been damaged by herbivorous insects, they react by transforming compounds in the odour into effective anti-herbivore defences.

When insects feed on plant tissues, the assaulted plant can exhibit a range of physiological responses. For example, it may produce chemical toxins and feeding deterrents, or emit airborne volatile compounds that attract natural enemies of the feeding herbivores, such as insect predators and parasitoids (for instance, parasitic wasps, which lay their eggs in plant-feeding caterpillars)<sup>2</sup>. It has become widely accepted<sup>3–6</sup> that plants can also use volatile emissions released by damaged neighbours as cues to prepare their own

defences against an impending attack — an idea that had previously been controversial<sup>3,4,7</sup>.

Sugimoto *et al.* investigated the mechanisms by which volatile signalling between cultivated tomato plants influences their defence against larvae of the moth *Spodoptera litura*, an agricultural pest also known as the common cutworm. Using an experimental set-up in which airflow between individual plants was carefully controlled, the authors showed that exposure to the odours released by cutworm-damaged tomato plants significantly enhanced the ability of neighbouring plants to resist a subsequent attack. Cutworm larvae placed on plants that had been exposed for three days to the odours released by damaged plants showed both reduced growth and increased mortality compared with larvae placed on unexposed control plants.

The authors' extensive biochemical analyses of tomato-leaf tissues revealed that plants exposed to volatiles that had been released by damaged neighbours had greatly elevated levels of a single compound — (Z)-3-hexenyl-vicianoside, or HexVic. Furthermore, they found that cutworm larvae fed an artificial diet laced with HexVic showed significantly reduced growth compared with larvae reared on an untainted diet, confirming that



**Figure 1 | Plant odours as alarm signals.** When plants are attacked by herbivorous insects, they release volatile compounds that can attract insect predators and deter further herbivory. These volatile odours also enhance the defences of neighbouring plants against attack. Sugimoto *et al.*<sup>1</sup> report that tomato plants can directly transform (Z)-3-hexenol, a volatile compound released by their damaged neighbours, into (Z)-3-hexenylvicianoside (HexVic), an effective defence compound that reduces the growth and survival of the herbivores. (Figure adapted from drawings by Nick Sloff and Thomas Degen.)

accumulation of HexVic is involved in the tomato plants' increased resistance. Sugimoto and colleagues were struck by the remarkable chemical similarity between HexVic and (Z)-3-hexenol, a component of the odour emitted by damaged tomato plants. This suggested that the transmission of (Z)-3-hexenol from damaged to undamaged plants might provide the raw material for HexVic production, rather than that the volatile was acting simply as a chemical signal (Fig. 1).

To test this, the authors exposed tomato plants to airborne (Z)-3-hexenol tagged with the hydrogen isotope deuterium, and found that all of the HexVic subsequently produced was labelled with the isotope. Furthermore, they found no indication that exposure to volatiles caused the plants to produce any extra (Z)-3-hexenol. It seems, therefore, that the accumulation of HexVic in undamaged tomato plants is entirely dependent on the uptake of (Z)-3-hexenol from the atmosphere. However, once plants are attacked by herbivores, their own (Z)-3-hexenol production seems to drive further accumulation of HexVic.

The remarkable mechanism of defence induction documented by Sugimoto *et al.* may be widespread in plants. (Z)-3-Hexenol is often found in 'green-leaf' volatiles, which most plants emit in the immediate aftermath of tissue damage (this compound contributes to the characteristic odour of freshly cut grass), and the biochemical transformation of (Z)-3-hexenol to HexVic is relatively simple and probably quite common. Indeed, the authors present evidence that HexVic, or chemically similar compounds, accumulate in many plant species after exposure to green-leaf volatiles.

This mechanism may also illuminate a plausible scenario for the evolution of plant-to-plant communication. From an evolutionary perspective, it is difficult to understand the benefit of signals that send reliable warnings of impending herbivory to neighbouring plants that will frequently be competitors of the emitter. Consequently, previous work<sup>6,8</sup> has suggested that plant-to-plant signalling might emerge as a by-product of volatile-signalling systems in individual plants. Consistent with this hypothesis, it is easy to envisage the mechanisms described by Sugimoto and colleagues as having initially evolved to regulate the accumulation of plant-defence compounds in distant parts of the emitting plant.

Even if neighbouring plants are not the intended receivers of green-leaf volatiles, they probably benefit from this intrinsic ability to produce HexVic from atmospheric (Z)-3-hexenol, because the levels of this compound in the air are likely to reflect the intensity of local herbivore activity. This effect might well have broader ecological implications, because an increase in volatile-induced resistance in local plant communities that

grows in proportion to the size and intensity of emerging insect infestations could act as a brake on the spread of the infestation.

Despite the far-reaching implications of this study, it is unlikely to be the last word on plant-to-plant signalling or plant olfaction. For example, Sugimoto *et al.* found no link between the mechanisms they observed and the expression of jasmonic acid (a plant hormone that regulates many defence mechanisms activated by herbivory<sup>9</sup>), although priming of jasmonic-acid-mediated defences by volatiles has been reported in different species. Moreover, other classes of plant compounds have been implicated in plant-plant interactions<sup>4,10</sup>, and a study last year<sup>11</sup> also reported plant-defence responses to insect odours that bore little or no resemblance to the typical emissions of plants.

It is unclear whether the mechanisms underlying other plant responses to environmental odours will prove as straightforward as the uptake and subsequent conversion of (Z)-3-hexenol to HexVic in tomato plants, or to what extent they parallel the far greater complexity observed in animal olfactory systems<sup>12</sup>. But it is certain that many more fascinating and

unexpected facets of plant olfaction remain to be discovered. ■

**Mark C. Mescher and Consuelo M. De Moraes** are in the Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland.  
e-mails: mescher@usys.ethz.ch;  
consuelo.demoraes@usys.ethz.ch

1. Sugimoto, K. *et al.* *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1320660111> (2014).
2. De Moraes, C. M., Lewis, W. J., Paré, P. W., Alborn, H. T. & Tumlinson, J. H. *Nature* **393**, 570–573 (1998).
3. Karban, R., Yang, L. H. & Edwards, K. F. *Ecol. Lett.* **17**, 44–52 (2014).
4. Farmer, E. E. *Nature* **411**, 854–856 (2001).
5. Karban, R., Baldwin, I. T., Baxter, K. J., Laue, G. & Felton, G. W. *Oecologia* **125**, 66–71 (2000).
6. Heil, M. & Silva Bueno, J. C. *Proc. Natl Acad. Sci. USA* **104**, 5467–5472 (2007).
7. Fowler, S. V. & Lawton, J. H. *Am. Nat.* **126**, 181–195 (1985).
8. Frost, C. J. *et al.* *Ecol. Lett.* **10**, 490–498 (2007).
9. Gfeller, A., Dubugnon, L., Liechti, R. & Farmer, E. E. *Sci. Signal.* **3**, cm3 (2010).
10. Runyon, J. B., Mescher, M. C. & De Moraes, C. M. *Science* **313**, 1964–1967 (2006).
11. Helms, A. M., De Moraes, C. M., Tooker, J. F. & Mescher, M. C. *Proc. Natl Acad. Sci. USA* **110**, 199–204 (2013).
12. Bushdid, C., Magnasco, M. O., Vosshall, L. B. & Keller, A. *Science* **343**, 1370–1372 (2014).

#### IMMUNOLOGY

## When lymphocytes run out of steam

**The finding that absence of the enzyme CTPS1 underlies a form of human immunodeficiency highlights the role of metabolism in immune responses and suggests avenues for treating diseases such as leukaemia. SEE LETTER P.288**

**ANDRÉ VEILLETTE & DOMINIQUE DAVIDSON**

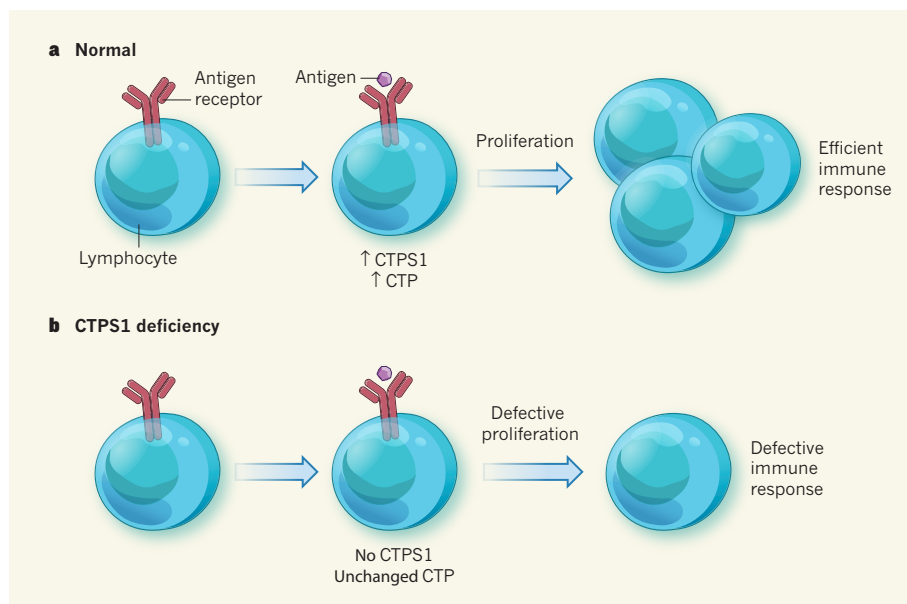
**L**ymphocytes are immune cells that play a crucial part in protection against micro-organisms such as viruses and bacteria. The two main types of lymphocyte are T cells and B cells, which, in the presence of antigen molecules derived from microbes, undergo a series of molecular changes that induce a state of activation. This response is driven by antigen receptors on the cells' surface, and leads to rapid cell proliferation and augmented immune protection (Fig. 1a). Proliferation under these conditions depends on metabolic adaptation, which allows immune cells to synthesize DNA and RNA molecules and the proteins needed for cell division<sup>1</sup>. In this issue, Martin *et al.*<sup>2</sup> (page 288) link this metabolic requirement of proliferating lymphocytes to a newly described immune-deficiency disease.

The authors describe children from several unrelated families who developed a severe immunodeficiency at birth or at a very

young age. In some cases, two members of the same family had the deficiency. Typically, the patients exhibited severe and persistent infections with viruses such as Epstein–Barr, the cause of infectious mononucleosis, and varicella zoster, which causes chickenpox and herpes zoster (shingles). Severe infections from bacteria such as pneumococcus, a cause of pneumonia, were also noted. Several children underwent transplantation with haematopoietic stem cells (which can differentiate into all types of blood cell) to control the infections. Immunological investigations led the authors to propose that the patients might be suffering from an inherited immunodeficiency that compromises lymphocyte function.

Sequencing of DNA from the affected children revealed that they all carried a mutation in *CTPS1*, the gene encoding the enzyme cytidine nucleotide triphosphate synthase 1 (CTPS1), that resulted in an absence of this enzyme in the patients' lymphocytes. CTPS1 is one of two forms of CTP synthase enzymes





**Figure 1 | Regulation of lymphocyte activation by CTPS1.** **a**, In normal lymphocytes (T and B cells), stimulation of the cells' antigen receptor triggers a series of molecular changes that induce the cells to proliferate, fuelling the immune response. Martin *et al.*<sup>2</sup> show that these events include an increase in levels of the enzyme CTPS1 and its product, CTP, which supports the increased DNA synthesis required for cell proliferation. **b**, In lymphocytes from CTPS1-deficient humans, stimulation by antigen results in some, but not all, of the molecular changes associated with lymphocyte activation. In particular, there is no increase in CTP levels in the activated cells, resulting in compromised DNA synthesis, reduced lymphocyte proliferation and an impaired immune response.

produced in mammalian cells (the other being CTPS2); both enzymes enable the production of cytidine nucleotide triphosphate (CTP), a nucleotide required for cellular DNA and RNA synthesis<sup>3</sup>.

The authors show that normal lymphocytes express both CTPS1 and CTPS2: CTPS1 is present at low levels before lymphocyte activation and becomes markedly expressed in activated lymphocytes, whereas CTPS2 is already expressed at high levels in non-activated lymphocytes. Analyses of T and B cells from the CTPS1-deficient patients revealed that differentiation of the cells in the absence of activation by foreign microorganisms was largely unaffected by the mutation, and the immediate molecular changes triggered by antigen-receptor stimulation were mostly unaltered. However, the cells' capacity to synthesize DNA and proliferate following stimulation of the antigen receptor was severely compromised. Intracellular levels of CTP were also very low. These defects were reproduced when CTPS1 expression was artificially reduced in normal lymphocytes, or when 3-deazauridine, a pharmacological inhibitor of CTPS enzymes, was used to suppress their activity. Conversely, the defects were corrected when CTPS1 was introduced into cells of CTPS1-deficient patients by retrovirus-mediated gene transfer, or when CTP was added to the cells' culture medium.

These findings show that CTPS1 and its product, CTP, are required for lymphocytes to proliferate intensely during antigen-induced activation, further highlighting the

importance of rapid metabolic adaptation for proper immunity. In the absence of CTPS1, antigen-stimulated lymphocytes do not produce sufficient quantities of CTP, causing defects in DNA synthesis and cell proliferation (Fig. 1b). These effects explain in large part why CTPS1-deficient children develop life-threatening viral and bacterial infections.

In addition to identifying the genetic cause of a new immunodeficiency, Martin and colleagues' results raise several prospects for future investigation. They indicate that even though CTPS2 is expressed in lymphocytes, it cannot replace CTPS1 when the latter is deficient. One possible explanation for this is that CTPS1 is much more active than CTPS2 in lymphocytes, possibly owing to differences in intrinsic activity or abundance, or to modifications such as phosphorylation or co-factor binding that could influence the enzymes' activity. For instance, in mammalian cells, CTPS1, but not CTPS2, can be regulated by phosphorylation on certain amino-acid residues<sup>4,5</sup>. Alternatively, because it has been reported that CTPS1 and CTPS2 can be localized in both the cytoplasm and the nucleus<sup>6</sup>, it is possible that CTPS1 accumulates in a locale that is especially important for generating the CTP needed for intense DNA synthesis. Moreover, these enzymes form tetramers and can exist as large filamentous structures<sup>7,8</sup>; whether differences in these arrangements exist between CTPS1 and CTPS2 remains to be clarified.

Most studies of CTPS enzymes have focused

on their capacity to promote DNA and RNA synthesis. However, investigations of the two yeast homologues of CTPS enzymes, URA7 and URA8, indicated that these enzymes are also needed for the synthesis of phospholipids<sup>9</sup>. It is possible that the use of CTP during the synthesis of membrane phospholipids is needed for the interaction of signalling molecules with the inner leaflet of the cell membrane. If such an activity exists for mammalian CTPS enzymes, defects in membrane-driven signalling could contribute to the lymphocyte dysfunctions observed in CTPS1-deficient patients. This activity could also explain Martin and colleagues' observation that CTPS1-deficient human T cells have reduced activation of the enzyme Erk kinase and reduced expression of the signal-transmission proteins CD25 and CD69 following antigen-receptor stimulation. These molecular events occur at early stages in cell activation, before the initiation of DNA synthesis, and are crucial for productive lymphocyte activation.

The data also raise the provocative possibility that pharmacological inhibitors of CTPS1 could be useful tools for treating human diseases associated with excessive or uncontrolled lymphocyte proliferation, such as transplant rejection, graft-versus-host disease and some forms of cancers such as leukaemia and lymphoma. In keeping with the latter idea, the CTPS inhibitor 3-deazauridine has already been shown to display some therapeutic efficacy against leukaemic cells *in vitro*, although it probably also inhibited targets other than CTPS in these cells<sup>10</sup>. The development of more-specific inhibitors of CTPS1 will aid the further investigation of this possible therapeutic avenue. ■

**André Veillette and Dominique Davidson** are in the Laboratory of Molecular Oncology, Institut de recherches cliniques de Montréal, Montreal, Quebec H2W1R7, Canada. A.V. is also in the Department of Medicine, University of Montreal and the Department of Medicine, McGill University, Montreal. e-mail: veillea@ircm.qc.ca

- Pearce, E. L., Poffenberger, M. C., Chang, C.-H. & Jones, R. G. *Science* **342**, 1242454 (2013).
- Martin, E. *et al.* *Nature* **510**, 288–292 (2014).
- Evans, D. R. & Guy, H. I. *J. Biol. Chem.* **279**, 33035–33038 (2004).
- Higgins, M. J., Graves, P. R. & Graves, L. M. *J. Biol. Chem.* **282**, 29493–29503 (2007).
- Kassel, K. M., Au, D. R., Higgins, M. J., Hines, M. & Graves, L. M. *J. Biol. Chem.* **285**, 33727–33736 (2010).
- Gou, K.-M., Chang, C.-C., Shen, Q.-J., Sung, L.-Y. & Liu, J.-L. *Exp. Cell Res.* **323**, 242–253 (2014).
- Levitzi, A. & Koshland, D. E. Jr *Biochemistry* **11**, 247–253 (1972).
- Liu, J.-L. *BioEssays* **33**, 159–164 (2011).
- Chang, Y.-F. & Carman, G. M. *Prog. Lipid Res.* **47**, 333–339 (2008).
- McPartland, R. P., Wang, M. C., Bloch, A. & Weinfeld, H. *Cancer Res.* **34**, 3107–3111 (1974).

This article was published online on 28 May 2014.

# The hunt for Majorana neutrinos hots up

**Finding that neutrinos are their own antiparticles would revolutionize particle physics. A high-sensitivity technique accelerates the search for the nuclear-decay process that would enable such a discovery. [SEE ARTICLE P.229](#)**

DAVID WARK

Discovering a new class of fundamental particle is about the biggest bang you can make in particle physics. The discovery of the Higgs boson was so exciting partly because it is a fundamental particle with no spin, the first ever seen. Physicists have long tried to resolve whether the familiar neutrino actually belongs to a class of exotic entities called Majorana particles, which are their own antiparticles. Majorana neutrinos might also help to explain why neutrinos are so light, and could be a clue to how the Universe came to contain so much more matter than antimatter. Almost the only practical way to reveal Majorana neutrinos would be to observe the rare nuclear process called neutrinoless double- $\beta$  decay. On page 229 of this issue, the EXO-200 Collaboration<sup>1</sup> announces the result of a high-sensitivity technique to search for this decay. Their results show the power of their technique, but demonstrate that there is still much work to do in the search for Majorana neutrinos.

Double- $\beta$  decay, as the name implies, occurs when a nucleus undergoes two  $\beta$  decays simultaneously and so emits two  $\beta$ -particles (electrons or antielectrons). This is realistically observable only in the few nuclei for which single- $\beta$  decay would lead to a heavier daughter nucleus but two  $\beta$  decays would lead to a lighter one. The standard model of particle physics allows two-neutrino double- $\beta$  decay ( $2\nu\beta\beta$  decay), which is just two ordinary  $\beta$  decays occurring at the same time, leading to the emission of two electrons and two antineutrinos. But what if, as many theorists believe, neutrinos are Majorana? Then another type of decay could also occur, in which the antineutrino emitted from one of the  $\beta$  decays could be absorbed as a neutrino in the other  $\beta$  decay, and the process seen from outside the nucleus would just be the emission of the two electrons, or zero-neutrino double- $\beta$  decay ( $0\nu\beta\beta$  decay).

$0\nu\beta\beta$  decay has a beautiful experimental signature — the simultaneous emission of two electrons with a total energy that sums to the difference in mass between the parent and daughter nuclei. Furthermore, it can happen only if neutrinos have mass.

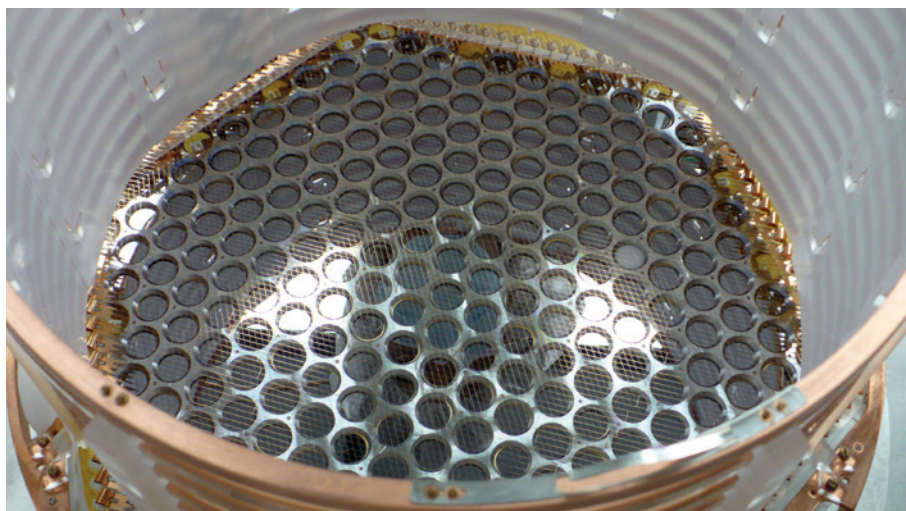
Owing to the nature of the weak interactions that cause  $\beta$  decay, the emitted antineutrino will have a spin that is right-handed with respect to its direction of motion, and the absorbed neutrino must be left-handed. This coexistence of the different spin states happens only for particles that have mass, and therefore makes the observed rate of  $0\nu\beta\beta$  decay a sensitive probe of the mass of the neutrino.  $0\nu\beta\beta$  decay is therefore a ‘two for the price of one’ measurement, which is one reason that so many groups are attacking it with different techniques and different nuclei.

The EXO (Enriched Xenon Observatory)-200 team has now moved near to the front of the pack with its liquid-xenon ‘time-projection chamber’ detector (Fig. 1). The active part of the detector is 110 kilograms of liquid xenon (which acts both as the double- $\beta$ -decay source and the detector) enriched to 80.6% in the potentially double- $\beta$ -decaying nucleus xenon-136 ( $^{136}\text{Xe}$ ). Charged particles — either the electrons emitted in double- $\beta$  decay or, much less desirably, background events from contaminant nuclei in the detector or surroundings — produce ionization in the xenon, and are drifted in an applied electric field to two

crossed grids of wires. The grids determine the position of the charges in two dimensions. The charged particles also make the xenon emit scintillation light, and measuring the time difference between that light and the arrival of the charge at the wires gives the drift distance (hence the name time-projection chamber), allowing reconstruction of the tracks from the charged particles in three dimensions.

This reconstruction is crucial because, as in all double- $\beta$ -decay experiments, the main challenge in EXO-200 is suppressing the background events. The first step towards this was to minimize the background from cosmic rays by locating the detector deep underground in the Waste Isolation Pilot Plant site near Carlsbad, New Mexico. More-insidious background events arise from radioactive contaminants in and around the detector, so all materials that make up the detector were carefully selected to have the lowest possible levels of radioactive contamination, and the detector itself is heavily shielded. Background events still persist, however, so another advantage of the time-projection chamber comes to the fore. By looking for events in which almost all of the charge is deposited at a single place in the detector (as one would expect for a double- $\beta$  decay) and rejecting those in which it comes from more than one place (as would be expected for  $\gamma$ -rays from radioactive contaminants), the background is suppressed by another large factor. Having these separate sets of data also helps to validate the properties of the detector, which is calibrated with an extensive set of measurements using radioactive sources.

The EXO-200 team has now reported the results of two years of running its experiment. The main output of the detector is a spectrum showing how often a particular energy was deposited within it.  $2\nu\beta\beta$  decay produces a



**Figure 1 | Chasing Majorana neutrinos.** The EXO-200 Collaboration<sup>1</sup> has used a liquid-xenon ‘time-projection chamber’ detector, half of which is seen here under construction, to search for the neutrinoless double- $\beta$ -decay process that would arise if neutrinos are Majorana-type particles. The two crossed grids of wires that were used to determine the position of charged particles in the detector, which is roughly 40 cm in diameter, lie above the frame of empty circular holes that were later filled with scintillation-light detectors.

EXO-200 COLLABORATION



## CANCER BIOLOGY

continuous energy spectrum that dominates the low-energy part of the observed range (see Fig. 4 of the paper<sup>1</sup>), showing the success of the background suppression. However, no statistically significant sign of a concentration of energy deposits around the energy expected from  $0\nu\beta\beta$  is apparent. The authors conclude that the half-life for  $0\nu\beta\beta$  decay is greater than  $1.1 \times 10^{25}$  years (at the 90% confidence level). Long lifetimes correspond to a low probability that the neutrino will flip spin, and so indicate small neutrino masses.

These results are bad news for a previous claim<sup>2</sup> of a positive signal for  $0\nu\beta\beta$  in the decay of germanium. However, that claim had already been put under strain by earlier results<sup>3</sup> from EXO-200, by KamLAND-Zen (another Xe experiment in Japan based on a different technique)<sup>4</sup> and by the GERDA experiment<sup>5</sup> on double- $\beta$  decay in germanium. In fact, the earlier results of EXO-200 had set the tighter lower limit on the  $0\nu\beta\beta$  half-life of  $1.6 \times 10^{25}$  years, and KamLAND-Zen a limit of  $1.9 \times 10^{25}$  years. But in both those cases the limit was better than expected because the number of  $0\nu\beta\beta$  candidate events seen was actually smaller than expected on the basis of the background events alone.

The new EXO-200 result, which is based on 3.8 times more data than its first result<sup>3</sup>, actually has a poorer limit, because now the researchers see a 1.2-standard-deviation excess over background in the region of the  $0\nu\beta\beta$ -decay energy line — that is, they see slightly more events than their background model would predict. Such an excess of events would show up almost once every six times just from statistical fluctuations. However, as we saw with the search for the Higgs boson, a real signal would show up initially as a statistically insignificant excess. The question is whether the excess will grow with further data, and for that we must wait for more results from all the experiments. Understanding the nature of neutrinos is pivotal to models of particle physics and cosmology, so, especially now that experiments on neutrino oscillations<sup>6</sup> have shown that neutrinos have mass, the search must continue. ■

**David Wark** is in the Department of Physics, University of Oxford, and in the Particle Physics Department, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK. e-mail: david.wark@stfc.ac.uk

1. The EXO-200 Collaboration. *Nature* **510**, 229–234 (2014).
2. Klapdor-Kleingrothaus, H. V. & Krivosheina, I. V. *Mod. Phys. Lett. A* **21**, 1547–1566 (2006).
3. Auger, M. *et al.* (EXO-200 Collaboration) *Phys. Rev. Lett.* **109**, 032505 (2012).
4. Gando, A. *et al.* (KamLAND-Zen Collaboration) *Phys. Rev. Lett.* **110**, 062502 (2013).
5. Agostini, M. *et al.* (GERDA Collaboration) *Phys. Rev. Lett.* **111**, 122503 (2013).
6. Fogli, G. L. *et al.* *Nucl. Phys. B Proc. Suppl.* **235–236**, 125–132 (2013).

This article was published online on 4 June 2014.

# Enzyme meets a surprise target

**An enzyme previously implicated in gene regulation has now been found to have a role in cancer progression, potentiating an intracellular signalling pathway that is driven by a mutated K-Ras protein. SEE LETTER P.283**

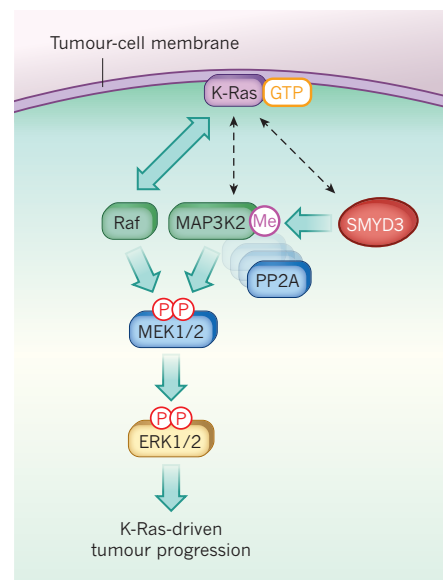
MARIAN M. DEUKER & MARTIN MCMAHON

Activating mutations in the gene *KRAS*, which encodes a member of the Ras protein family, are implicated in the development of many human cancers<sup>1</sup>. However, because drugs that effectively treat these cancers by targeting the K-Ras protein have proved difficult to develop, the search for potential therapeutic targets has turned to the proteins that are activated downstream of this oncoprotein. In this issue, Mazur *et al.*<sup>2</sup> (page 283) identify the enzyme SMYD3 as a protein that has an unanticipated role in the progression of K-Ras-driven cancers in mice.

SMYD3, a lysine methyltransferase, is frequently overexpressed in human cancers<sup>2,3</sup>. Previous work<sup>3</sup> has indicated that the enzyme primarily acts in the nucleus, adding methyl groups to lysine amino-acid residues on histones (proteins that organize DNA into bundles called nucleosomes). To investigate the role of SMYD3 in K-Ras-driven cancers, Mazur and colleagues used several techniques, spanning the gamut from genetically engineered mouse models to screens for SMYD3 protein substrates.

The authors found that in mouse models of K-Ras-driven cancer, SMYD3 acts in the cytoplasm of cancer cells, methylating a lysine residue (K260) on MAP3K2, a kinase enzyme that is associated<sup>4</sup> with the activation of several stress-induced pathways. They report that methylation of MAP3K2 potentiates a cellular signalling pathway that is involved in many human cancers<sup>5</sup>, the MEK–ERK mitogen-activated protein-kinase pathway. This pathway, which activates gene transcription in the nucleus, is composed of a cascade of enzymes that are sequentially activated in the presence of oncogenic K-Ras. Typically, K-Ras activates Raf; Raf activates MEK; and MEK activates ERK. Mazur and colleagues' work demonstrates for the first time that MAP3K2 can also potentiate MEK–ERK signalling downstream of oncogenic K-Ras (Fig. 1).

The researchers show that SMYD3-mediated methylation of MAP3K2 has no direct effect on the protein's intrinsic enzymatic activity. Instead, methylation promotes dissociation of MAP3K2 from its negative regulator, the PP2A phosphatase enzyme complex. Consistent



**Figure 1 | SMYD3 potentiates signalling in K-Ras-driven cancer.** Mutational activation of *KRAS* causes increased signalling through the MEK–ERK pathway, leading to cancer. In this pathway, K-Ras (which is bound by a GTP molecule) activates the kinase enzyme Raf, leading to a kinase cascade in which MEK1 and MEK2 and then ERK1 and ERK2 are activated by phosphorylation (P). ERK1 and ERK2 then activate genes involved in K-Ras-driven tumour progression. Mazur *et al.* report<sup>2</sup> that the lysine methyltransferase enzyme SMYD3 is involved in K-Ras-driven tumour progression. The enzyme acts by adding methyl groups (Me) to a lysine amino-acid residue of the kinase MAP3K2, causing it to dissociate from its negative regulator PP2A, and further activating the MEK–ERK pathway. However, it remains unclear precisely how SMYD3 and MAP3K2 are linked to K-Ras (dashed arrows).

with a role for SMYD3 in tumour progression, Mazur and co-workers found that silencing of SMYD3 lengthens the median lifespan of mice that have been genetically engineered to express oncogenic K-Ras in the lungs or pancreas. Interestingly, they saw this effect only in mice with late-stage cancer progression; SMYD3 silencing did not influence tumour initiation. Strikingly, administration of a PP2A inhibitor restored tumour progression in the SMYD3-deficient mice. However, the authors did not test the effects of SMYD3 silencing in mice with aggressive cancers, which would

more accurately model the human disease.

In sum, Mazur *et al.* propose that elevated SMYD3 expression promotes MAP3K2 methylation, freeing the protein from the inhibitory constraints of PP2A and thus potentiating signalling downstream of oncogenic K-Ras. Interestingly, methylation has previously been shown<sup>6</sup> to play a part in the regulation of Raf–MEK–ERK signalling — methylation of the Raf proteins B-Raf or C-Raf promotes their degradation, thereby inhibiting MEK–ERK signalling. The fact that protein methylation has now emerged as both a positive and a negative regulator of MEK–ERK-pathway activation suggests that this might be a common strategy for the regulation of intracellular signalling pathways.

Mazur and colleagues' work raises the question of how SMYD3 links MAP3K2 and oncogenic K-Ras. It is unclear why SMYD3 levels are elevated in Ras-mutated cancers, such as colorectal<sup>7</sup> and pancreatic cancer<sup>2</sup>, although it is possible that Ras regulates SMYD3 expression through effects on either gene transcription or protein stability. Despite early claims<sup>7</sup> that MAP3K2's sister enzyme MAP3K1 can bind directly to activated Ras proteins, it seems unlikely that MAP3K2 interacts directly with oncogenic K-Ras.

Although it is well established that MEK–ERK signalling is involved in K-Ras-driven cancer, this pathway's dependence on different MAP3K enzymes (B-Raf, C-Raf and MAP3K2) is influenced by both the stage of the cancer and the tissue from which it originated<sup>2,8–10</sup>. Thus, the differing effects of SMYD3 silencing on early- and late-stage cancer progression may arise because of differing levels of MEK–ERK-pathway activation. Perhaps the activation of MAP3K2 is tied to a signalling circuit that, although not essential for the activation of MEK–ERK signalling in early-stage cancers, is required for the elevated activation of this pathway in more-advanced tumours<sup>11</sup>. Given the number of growth factors and other signalling molecules whose secretion is reported to be elevated in K-Ras-driven cancers<sup>12–16</sup>, oncogenic K-Ras may promote MAP3K2 activation through a feed-forward loop driven by intra- or intercellular signalling proteins, such as members of the epidermal-growth-factor family. The relationship between oncogenic K-Ras and SMYD3 is intricate and remains unclear. Moreover, at least one complication with this model is the observation that the protein kinase MKK7, a direct downstream target of MAP3K2, suppresses K-Ras-initiated lung-tumour progression<sup>17</sup>.

Might patients with late-stage K-Ras-driven lung or pancreatic cancer benefit from drugs that target SMYD3 or MAP3K2? Mazur *et al.* demonstrate that SMYD3 silencing slows the proliferation of human K-Ras-mutated cancer cells. Furthermore, they found that silencing of SMYD3 enhanced the antitumour effects of trametinib, a MEK–ERK-pathway inhibitor, in

both cultured cells and genetically engineered mice. Importantly, mice lacking SMYD3 or MAP3K2 seem normal<sup>2,18</sup>, suggesting that inhibitors of these proteins do not have severe side effects, at least in mice.

Efforts to develop drugs that inhibit SMYD-family lysine methyltransferases or MAP3K2 are already under way<sup>19,20</sup>. However, inhibition of MAP3K2 alone may not be an effective way to dampen MEK–ERK signalling because other MAP3K enzymes might be able to perform the role of MAP3K2 in its absence. Combined inhibition of SMYD3 and MAP3K2 may therefore produce more promising results.

Despite Mazur and colleagues' claims about the specificity of SMYD3 for MAP3K2, previous work<sup>3</sup> suggests that SMYD3 also acts as a histone modifier, regulating RNA polymerase II, the enzyme responsible for catalysing gene transcription. Thus, inhibition of SMYD3 might cause unexpected side effects. Nevertheless, the authors' work has expanded our understanding of the regulatory role of protein methylation in intracellular signalling. We must wait to see whether this regulatory mechanism can be manipulated to treat patients with K-Ras-mutated cancers. ■

**Marian M. Deuker and Martin McMahon**  
*are in the Helen Diller Family Comprehensive Cancer Center and the Department of Cell and Molecular Pharmacology, University of California, San Francisco,*

*San Francisco, California 94158, USA.*  
*e-mail: mcmahon@cc.ucsf.edu.*

1. Pylayeva-Gupta, Y., Grabocka, E. & Bar-Sagi, D. *Nature Rev. Cancer* **11**, 761–774 (2011).
2. Mazur, P. K. *et al. Nature* **510**, 283–287 (2014).
3. Hamamoto, R. *et al. Nature Cell Biol.* **6**, 731–740 (2004).
4. Cuevas, B. D., Abell, A. N. & Johnson, G. L. *Oncogene* **26**, 3159–3171 (2007).
5. Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. *Oncogene* **26**, 3279–3290 (2007).
6. Andreu-Pérez, P. *et al. Sci. Signal.* **4**, ra58 (2011).
7. Russell, M., Lange-Carter, C. A. & Johnson, G. L. *J. Biol. Chem.* **270**, 11757–11760 (1995).
8. Karreth, F. A., Frese, K. K., DeNicola, G. M., Baccarini, M. & Tuveson, D. A. *Cancer Discov.* **1**, 128–136 (2011).
9. Blasco, R. B. *et al. Cancer Cell* **19**, 652–663 (2011).
10. Cronan, M. R. *et al. Oncogene* **31**, 3889–3900 (2012).
11. Feldser, D. M. *et al. Nature* **468**, 572–575 (2010).
12. McCarthy, S. A., Samuels, M. L., Pritchard, C. A., Abraham, J. A. & McMahon, M. *Genes Dev.* **9**, 1953–1964 (1995).
13. Schulze, A., Lehmann, K., Jefferies, H. B. J., McMahon, M. & Downward, J. *Genes Dev.* **15**, 981–994 (2001).
14. Sparrmann, A. & Bar-Sagi, D. *Cancer Cell* **6**, 447–458 (2004).
15. Ancrile, B., Lim, K.-H. & Counter, C. M. *Genes Dev.* **21**, 1714–1719 (2007).
16. Zhao, Q. & Lee, F. S. *J. Biol. Chem.* **274**, 8355–8358 (1999).
17. Schramek, D. *et al. Nature Genet.* **43**, 212–219 (2011).
18. Guo, Z. *et al. Mol. Cell. Biol.* **22**, 5761–5768 (2002).
19. Ferguson, A. D. *et al. Structure* **19**, 1262–1273 (2011).
20. Ahmad, S., Hughes, M. A., Johnson, G. L. & Scott, J. E. *J. Biomol. Screening* **18**, 388–399 (2013).

This article was published online on 21 May 2014.

## GENE THERAPY

# Repair and replace

**One approach to treating inherited diseases is repairing the defective genes, but this has proved challenging in stem cells. An optimized protocol has now been developed that allows gene repair in blood-cell precursors. [SEE ARTICLE P.235](#)**

ALAIN FISCHER

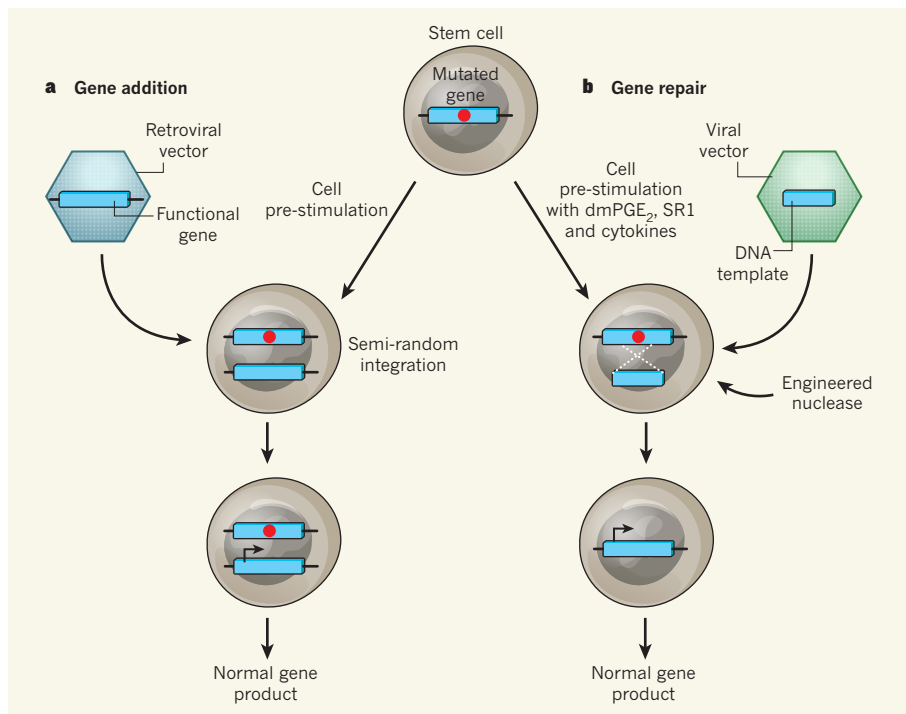
For more than 40 years<sup>1</sup>, gene therapy has been predicted to be a way to cure inherited diseases that are caused by defective copies of a gene. The typical approach to gene therapy is to add a functional copy of the mutated gene to the genome, using a genetically engineered retroviral vector to transport the copy into diseased cells (Fig. 1a). Although gene addition has been used to treat certain inherited diseases, such as severe combined immunodeficiencies<sup>2,3</sup>, its success has been mitigated by side effects<sup>4</sup>. A different strategy, based on repair of the defective gene, is therefore an attractive alternative. On page 235 of this issue, Genovese *et al.*<sup>5</sup> report progress in the optimization of such a strategy.

Gene-repair techniques rely on artificial nuclease enzymes that specifically target the

mutated genetic sequence and create a DNA break. If a DNA template for the desired replacement sequence is provided, the mutated stretch can be repaired by an exchange of genetic information known as homologous recombination (Fig. 1b)<sup>6</sup>. One benefit of this strategy is that it takes advantage of an innate method of recombination used by the cell to repair harmful DNA breaks, and so does not adversely affect other genomic regions, such as regulatory sequences. Gene addition, on the other hand, relies on a semi-random integration method that may disrupt genetic regulatory sequences, because the extra sequence integrates imprecisely within the genome. This has been shown<sup>4</sup> to cause toxic side effects, for example vector-mediated activation of cancer-causing oncogenes.

Zinc-finger nucleases (a zinc-finger protein that binds to the desired sequence of DNA,





**Figure 1 | Strategies for treating inherited diseases with gene therapy.** **a**, In gene addition, stem cells harbouring a genetic mutation are pre-stimulated to prepare them for genetic manipulation, and then a functional copy of the mutated gene is transported into the cells in a retroviral vector. The functional gene integrates into the genome at a semi-random site, restoring normal gene expression. **b**, An alternative approach is gene repair, whereby a nuclease enzyme engineered to create DNA breaks around the mutated site and a DNA template for the functional sequence are added to pre-stimulated cells using a viral vector. The mutated sequence is then replaced with the functional copy by homologous recombination (white dotted lines). Genovese *et al.*<sup>5</sup> have optimized pre-stimulation during gene repair in haematopoietic stem cells, by treating the cells both with factors that prevent early differentiation (dmPGE<sub>2</sub> and SR1) and with signalling molecules called cytokines that decrease the cells' sensitivity to the toxic effects of the nuclease.

combined with a nuclease, which cleaves DNA wherever the molecule binds) were the first enzymes to be designed for gene repair and have been shown to work in cell lines<sup>6</sup>. Other nucleases — for example, those modified from artificial restriction enzymes known as TALENs, or from RNA-guided enzymes based on the bacterial CRISPR-associated system — have since been used<sup>6</sup> to correct genetic mutations in human cell lines, including induced pluripotent stem cells (iPSCs), which have been genetically reset and can give rise to most cell types in the body. Genovese and colleagues have used this approach to correct mutations in a fraction of human haematopoietic stem cells (HSCs), which are a key target for treating inherited disorders associated with blood cells.

How did the authors achieve this? The hurdle to overcome lay in the fact that homologous recombination occurs only in cycling cells (in the S/G2 phase of the cell cycle), and so the procedure is ineffective in adult stem cells such as HSCs, which are mostly quiescent. Using several tricks, Genovese *et al.* achieved efficient integration of a green fluorescent protein sequence into the genome of human HSCs at chosen sites, including a protein-coding region (exon 5) of the *IL2RG* gene that is mutated in patients with X-linked severe combined immunodeficiency syndrome (SCID-X1).

To prepare HSCs for gene repair, Genovese and colleagues stimulated the cells with signalling molecules called cytokines for two days, and on the second day used a viral vector to introduce a recombination template. They then used electric currents to permeabilize the cells, and added a Zn-finger nuclease designed to target the desired DNA sequence. It is likely that pretreatment with cytokines made the HSCs less sensitive to the toxic effects of introducing the nucleases<sup>4</sup> and, more importantly, that the treatment prompted some of the cells to enter the cell cycle and become capable of mediating homologous recombination. In addition, the authors treated the cells with two compounds — dimethyl prostaglandin E<sub>2</sub> (dmPGE<sub>2</sub>) and SR1, an inhibitor of the aryl-hydrocarbon receptor protein — that prevent premature differentiation, keeping the HSCs in a stem-cell state.

Genovese and co-workers' protocol enabled site-specific genome editing in some of the HSCs. When the treated cells were transferred into immunodeficient mice, edited cells were detected for up to 18 weeks, and could even be transferred to a second recipient. When their technique was applied to HSCs taken from a patient with SCID-X1, the authors found that gene correction had occurred in 3–11% of progenitor cells and the myeloid cells that differentiated from them.

This work is undoubtedly a step towards using gene repair for gene therapy. Correction of SCID-X1 is known<sup>2</sup> to require only a few *IL2RG*-expressing HSCs, and so it may be that this technology can be used as it stands to treat patients with SCID-X1. But before that can be tested, it will be necessary to evaluate the frequency at which the protocol causes harmful DNA mutations, brought about by the cells' attempts to repair off-target DNA breaks through an alternative, error-prone mode of recombination. It is reassuring, however, that Genovese and co-workers observed few such off-target effects when they tested their protocol in a precancerous cell line.

If this method of gene repair in HSCs is to be applied to diseases that demand a higher level of gene correction than SCID-X1 (for example, Wiskott–Aldrich syndrome, adrenoleukodystrophy, metachromatic leukodystrophy or  $\beta$ -thalassaemia<sup>7–10</sup>), further optimization is required. For example, the design of nucleases that increase the fidelity of DNA cuts, combined with the *in vitro* expansion of treated HSCs before transfusion might increase the efficiency of the protocol. It is interesting to note that an improved gene-addition protocol, which uses self-inactivating viral vectors designed to reduce activation of host genes, including oncogenes, has led to the successful and apparently safe treatment of the inherited disorders mentioned above<sup>7–10</sup>. Competition between the two strategies is likely to continue for some time.

Ultimately, gene repair seems to be an attractive strategy for targeting stem cells, such as HSCs, that have been taken from tissues affected by inherited disorders. As technology that allows reprogramming of differentiated cells to stem-cell states improves, gene correction in reprogrammed cells such as iPSCs or induced HSCs<sup>11</sup> engineered from the patient's blood cells will probably become a preferable option. Until this goal is reached, however, it is good to see that Genovese and colleagues have made another advance in extending the usefulness of gene therapy. ■

**Alain Fischer** is at the Imagine Institute, Hôpital Necker-Enfants Malades, Paris 75015, France, and at the Collège de France, Paris. e-mail: alain.fischer@inserm.fr

1. Friedmann, T. & Roblin, R. *Science* **175**, 949–955 (1972).
2. Cavazzana-Calvo, M. *et al.* *Science* **288**, 669–672 (2000).
3. Aiuti, A. *et al.* *Science* **296**, 2410–2413 (2002).
4. Hacein-Bey-Abina, S. *et al.* *Science* **302**, 415–419 (2003).
5. Genovese, P. *et al.* *Nature* **510**, 235–240 (2014).
6. Hyongbum, K. & Jin-Soo, K. *Nature Rev. Genet.* **15**, 321–334 (2014).
7. Aiuti, A. *et al.* *Science* **341**, 1233151 (2013).
8. Cartier, N. *et al.* *Science* **326**, 818–823 (2009).
9. Biffi, A. *et al.* *Science* **341**, 1233158 (2013).
10. Cavazzana-Calvo, M. *et al.* *Nature* **467**, 318–322 (2010).
11. Riddell, J. *et al.* *Cell* **157**, 549–564 (2014).

This article was published online on 28 May 2014.

# Search for Majorana neutrinos with the first two years of EXO-200 data

The EXO-200 Collaboration\*

Many extensions of the standard model of particle physics suggest that neutrinos should be Majorana-type fermions—that is, that neutrinos are their own anti-particles—but this assumption is difficult to confirm. Observation of neutrinoless double- $\beta$  decay ( $0\nu\beta\beta$ ), a spontaneous transition that may occur in several candidate nuclei, would verify the Majorana nature of the neutrino and constrain the absolute scale of the neutrino mass spectrum. Recent searches carried out with  $^{76}\text{Ge}$  (the GERDA experiment) and  $^{136}\text{Xe}$  (the KamLAND-Zen and EXO (Enriched Xenon Observatory)-200 experiments) have established the lifetime of this decay to be longer than  $10^{25}$  years, corresponding to a limit on the neutrino mass of 0.2–0.4 electronvolts. Here we report new results from EXO-200 based on a large  $^{136}\text{Xe}$  exposure that represents an almost fourfold increase from our earlier published data sets. We have improved the detector resolution and revised the data analysis. The half-life sensitivity we obtain is  $1.9 \times 10^{25}$  years, an improvement by a factor of 2.7 on previous EXO-200 results. We find no statistically significant evidence for  $0\nu\beta\beta$  decay and set a half-life limit of  $1.1 \times 10^{25}$  years at the 90 per cent confidence level. The high sensitivity holds promise for further running of the EXO-200 detector and future  $0\nu\beta\beta$  decay searches with an improved Xe-based experiment, nEXO.

Majorana fermions, a class of neutral spin-1/2 particles described by two-component spinors, have been an element of quantum field theory since its inception<sup>1,2</sup>. Electrons and other spin-1/2 elementary particles with distinct antiparticles are, however, described by four-component Dirac spinors. Majorana quasiparticles may have been observed in condensed matter systems<sup>3</sup> where neutrality is achieved through the collective action of electrons and holes. Among the known elementary particles, only neutrinos are Majorana fermion candidates, owing to their intrinsic neutrality. Confirmation of this property would imply the non-conservation of lepton number, an additive quantum number that, unlike charge or colour, is not related to any known gauge symmetry. As yet, lepton number has been empirically found to be conserved. Neutrinos are also remarkable for their small, yet finite, masses<sup>4</sup> that are generally difficult to explain, but arise naturally in many extensions<sup>5,6</sup> of the standard model of particle physics. A generic consequence of many such extensions is that neutrinos should be of the Majorana variety.

The most sensitive probe for Majorana neutrinos is a nuclear process known as neutrinoless double- $\beta$  decay ( $0\nu\beta\beta$ ), whereby a nucleus decays by emitting two electrons and nothing else, while changing its charge by two units<sup>7</sup>. A related double- $\beta$  decay process, known as two-neutrino double- $\beta$  decay ( $2\nu\beta\beta$ ), is allowed by the standard model and has been observed in many nuclei,  $^{136}\text{Xe}$  among them<sup>8,9</sup>. It provides, however, no direct information on the Majorana/Dirac question. The exotic  $0\nu\beta\beta$  can be distinguished from the  $2\nu\beta\beta$  by measuring the sum energy of the two electrons that is peaked at the  $Q$ -value for the former and is a continuum for the latter (the  $Q$ -value is the mass difference between the mother and daughter nuclei). We refer to this region around the  $Q$ -value as the  $0\nu\beta\beta$  region of interest (ROI). The half-life of the  $0\nu\beta\beta$  is related to the effective Majorana neutrino mass ( $\langle m_{\beta\beta} \rangle$ ) by a phase space factor and a nuclear matrix element. Hence observation of the  $0\nu\beta\beta$  decay would discover elementary Majorana particles, demonstrate lepton number violation and measure the neutrino mass scale  $\langle m_{\beta\beta} \rangle$ , at least to within the theoretical uncertainty of the nuclear matrix elements<sup>10</sup>.

Recent sensitive searches for  $0\nu\beta\beta$  have been carried out in  $^{76}\text{Ge}$  (GERDA<sup>11</sup>) and  $^{136}\text{Xe}$  (KamLAND-Zen<sup>12</sup> and EXO-200<sup>13</sup>). These

experiments have set limits on the Majorana neutrino mass of  $\sim 0.2$ – $0.4$  eV, and have cast doubt on an earlier claim of observation<sup>14</sup>. In this Letter we report on new  $0\nu\beta\beta$  search results from the EXO-200 experiment based on about two years of data.

## The EXO-200 detector

EXO-200 has been described in detail elsewhere<sup>15</sup>. Briefly, the detector is a cylindrical liquid xenon (LXe) time projection chamber (TPC), roughly 40 cm in diameter and 44 cm in length. Two drift regions are separated in the centre by a cathode. The LXe is enriched to 80.6% in  $^{136}\text{Xe}$ , the  $0\nu\beta\beta$  candidate ( $Q = 2,457.83 \pm 0.37$  keV; ref. 16). The TPC provides  $X$ – $Y$ – $Z$  coordinate and energy measurements of ionization deposits in the LXe by simultaneously collecting the scintillation light and the charge. Charge deposits spatially separated by about 1 cm or more are individually observed and the position accuracy for isolated deposits is a few millimetres. Avalanche photodiodes (APDs) measure the scintillation light. Small radioactive sources can be positioned at standard positions near the TPC to calibrate the detector and monitor its stability.

The TPC is shielded from environmental radioactivity on all sides by  $\sim 50$  cm of HFE-7000 cryofluid<sup>17</sup> (HFE) maintained at  $\sim 167$  K inside a vacuum-insulated copper cryostat. Further shielding is provided by at least 25 cm of lead in all directions. The entire assembly is housed in a clean-room located underground at a depth of  $1,585^{+11}_{-6}$  metres water equivalent<sup>18</sup> (a measure of the effective shielding accounting for variations in the overhead rock) at the Waste Isolation Pilot Plant near Carlsbad (New Mexico). Four of the six sides of the clean-room are instrumented with plastic scintillator panels ('muon-veto panels') recording the passage of cosmic ray muons. An extensive materials screening campaign<sup>19</sup> was employed to minimize the radioactive background produced by the detector components.

## Data analysis and methodology

The data analysis methods in this work follow closely those presented in detail elsewhere<sup>9</sup>. Events in the detector are classified as single-site

\*A list of authors and affiliations appears at the end of the paper.



(SS) or multi-site (MS) according to the number of detected charge deposits.  $0\nu\beta\beta$  events are predominantly SS whereas  $\gamma$  backgrounds are mostly MS. For each event, the energy is determined as a linear combination of charge and scintillation, while a ‘standoff distance’ is defined as the distance between a charge deposit and the closest material that is not LXe, other than the cathode. To search for  $0\nu\beta\beta$ , a binned maximum-likelihood fit is performed simultaneously over the SS and MS events using probability density functions (PDFs) in energy and standoff distance, generated using a Geant4-based<sup>20</sup> Monte Carlo simulation (MC). The energy range 980–9,800 keV is used. The ‘low-background data set’ (physics data) is obtained after applying event selection cuts. With respect to ref. 9, the current analysis additionally includes: (1) improved signal processing for the scintillation waveforms resulting in lower noise; (2)  $^{226}\text{Ra}$  source calibration data; (3) an expanded fiducial volume; (4) the estimation of systematic errors related to the  $0\nu\beta\beta$  ROI; and (5) updated background and systematic studies relevant to the  $0\nu\beta\beta$  search.

The data set presented here (Run 2) combines Run 2a (already used for refs 9 and 13, 22 September 2011 to 15 April 2012) and Runs 2b and 2c (16 April 2012 to 1 September 2013). After removing periods of poor data quality and calibration runs, the total amount of low-background data for this analysis is  $477.60 \pm 0.01$  days, a 3.8-fold increase from previous EXO-200 publications. The primary tool used for understanding and correcting the detector energy measurement is the 2,615-keV  $\gamma$  line of  $^{208}\text{Tl}$  from a  $^{228}\text{Th}$  source deployed at least twice weekly during the time spanned by this data set. Seven multiday calibration campaigns involving the use of multiple sources ( $^{228}\text{Th}$ ,  $^{60}\text{Co}$ ,  $^{226}\text{Ra}$  and  $^{137}\text{Cs}$ ) were performed at roughly three-month intervals throughout the data set. The lifetime of ionization electrons in the LXe is better than 2 ms for the entire data set, more than sufficient to collect charge across the full volume of the detector. We determine the optimal linear combination of scintillation and ionization signals once per week by minimizing the width of the 2,615-keV line. To prevent making analysis decisions that could bias the results in the ROI, the low-background data were partially ‘masked’ to hide  $\sim 2/3$  of the live-time for SS events between 2,325 and 2,550 keV. Live-time already analysed in previous publications (for example, Run 2a) was not masked.

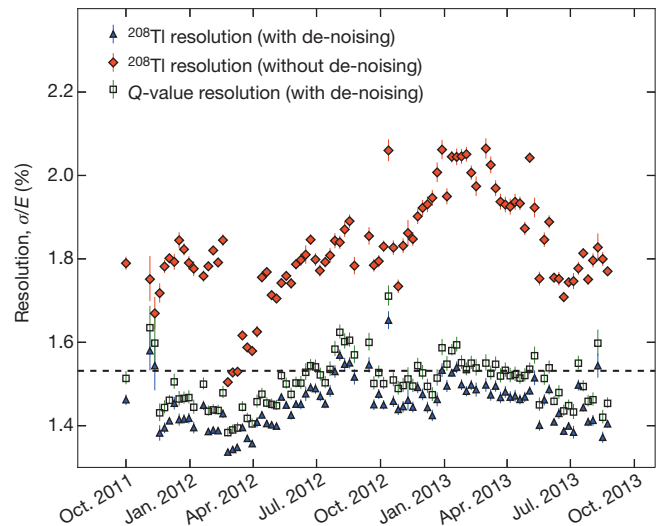
The energy resolution of the detector is dominated by electronic noise in the scintillation readout and exhibits variations over time due to changes in this noise. We apply a de-noising algorithm to the scintillation signals during post-processing, improving the detector resolution and reducing its time dependence. This algorithm attempts to find the optimal combination of APD waveforms to determine the amount of scintillation light for each event, taking into account the measured electronic noise of each APD channel as well as the position of each charge deposition in the detector. Figure 1 shows the resolution with and without de-noising.

We define an effective, time-independent energy resolution function<sup>9</sup>  $\sigma^2(E) = \sigma_{\text{elec}}^2 + b^2E + c^2E^2$ . Here  $\sigma_{\text{elec}}$ ,  $b$  and  $c$  are 20.8 keV,  $0.628 \text{ keV}^{1/2}$  and  $1.10 \times 10^{-3}$  (25.8 keV,  $0.602 \text{ keV}^{1/2}$  and  $4.04 \times 10^{-3}$ ) for SS (MS), determined by a maximum-likelihood fit to calibration data taken during Run 2. This function is folded with the energy distributions derived from the simulation to create the PDFs used in final fits. The effective resolution ( $\sigma/E$ ) for SS (MS) at the  $0\nu\beta\beta$  Q-value is  $1.53 \pm 0.06\%$  ( $1.65 \pm 0.05\%$ ).

The fiducial volume is larger than in ref. 9 to maximize the sensitive mass while maintaining systematic uncertainties at an acceptable level. Events in the fiducial volume are required to have  $182 \text{ mm} > |Z| > 10 \text{ mm}$  (where  $Z = 0$  is the cathode plane) and are contained in a hexagon with 162 mm apothem. This represents a  $^{136}\text{Xe}$  mass of 76.5 kg, corresponding to  $3.39 \times 10^{26}$  atoms of  $^{136}\text{Xe}$  and, with the quoted live-time, results in an exposure of 100 kg yr (736 mol yr).

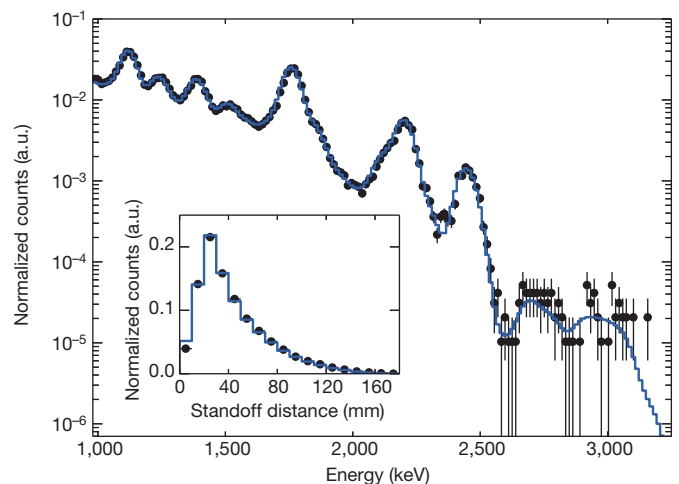
## Investigation and determination of systematic errors

The main systematic uncertainties relevant to the search for  $0\nu\beta\beta$  are related to signal efficiency, location of the  $0\nu\beta\beta$  ROI within the spectrum, and estimation of the background in the ROI.



**Figure 1 | Effect of de-noising on the energy resolution,  $\sigma/E$ .** Shown is the resolution for SS events at the 2,615-keV  $^{208}\text{Tl}$  full-absorption peak (with and without de-noising) and propagated to the  $0\nu\beta\beta$  Q-value (with de-noising). The variation with time (shown on x axis) is caused by changes in the noise of the APD front-end electronics. The horizontal dashed line shows the effective Q-value SS energy resolution used for the data set (1.53%). MS resolution (not shown) exhibits similar behaviour. Error bars,  $\pm 1$  s.d.

To verify the simulation’s ability to model efficiencies and the background, we compare measurement and simulation of calibration sources deployed at various positions around the TPC, investigating in particular: (1) the energy and standoff distance distributions, (2) the integrated rate of selected events, and (3) the SS/MS event ratio versus energy. A representative set of results for (1) is shown in Fig. 2, where simulation–data agreement for the  $^{226}\text{Ra}$  source are presented.  $^{226}\text{Ra}$  is a particularly valuable source because of several  $\gamma$  lines that map a broad energy region including the  $0\nu\beta\beta$  ROI. The energy spectrum shows good agreement across the energy range of the analysis. Comparable results were also obtained with the  $^{60}\text{Co}$  and  $^{228}\text{Th}$  sources. The stand-off distance agreement is within statistical errors except in the first 10 mm bin, where the simulation produces more events in the fiducial volume than seen in data.



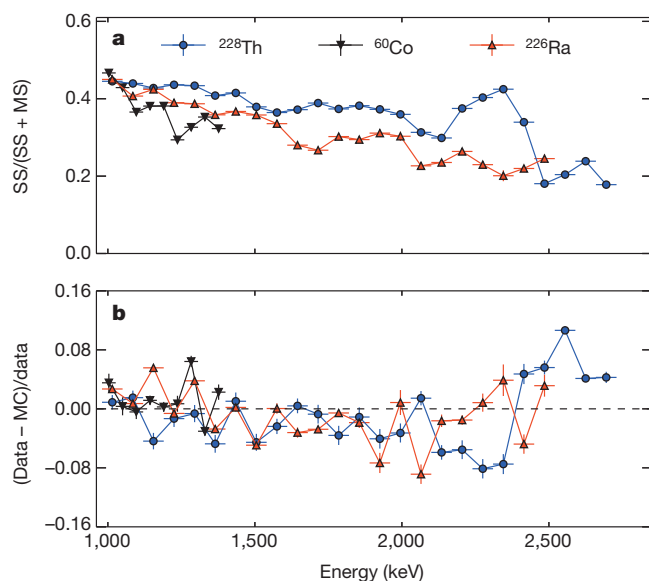
**Figure 2 | Comparison of energy and standoff distance distributions of a  $^{226}\text{Ra}$  calibration source for SS events in simulation and data.** Energy (main panel) and standoff distance (inset), both in normalized counts, are shown for data (black points) and simulation (blue line). The calibration source is at a position near the cathode outside the TPC. Error bars,  $\pm 1$  s.d.

Discrepancies in the shapes of energy and standoff distance distributions between data and simulation affect the estimation of the background in the  $0\nu\beta\beta$  ROI. To quantify this effect, we calculate skewing functions based on the small discrepancies observed in source calibration studies. We distort the background PDFs with the skewing functions and use these to produce a set of ‘toy’ MC data sets, which are then fitted to un-skewed PDFs. The change in the  $0\nu\beta\beta$  ROI background is 9.2%, which we take as systematic error.

In the rate comparison studies (2), we combine the total number of selected events in data and simulation as  $(\text{data} - \text{MC})/\text{data}$  for several source positions. The error-weighted average of the results is calculated using the fiducial volume in this analysis as well as that in ref. 9. The difference between these values is 1.7%, which we combine with the underlying fiducial volume uncertainty (also 1.7%; ref. 9) conservatively assuming full correlation to produce a total error on the detector efficiency of 3.4%.

To address (3), the ratio of the number of SS events to the total number of events,  $\text{SS}/(\text{SS} + \text{MS})$ , is compared between data and simulation for three sources in Fig. 3. The general behaviour is largely independent of the underlying spectral shape. We choose to assign a single systematic uncertainty to the  $\text{SS}/(\text{SS} + \text{MS})$  ratio of 9.6%, calculated from the weighted average of the maximum deviations observed for the  $^{228}\text{Th}$ ,  $^{60}\text{Co}$  and  $^{226}\text{Ra}$  (data from the latter available after June 2013) sources at several different source locations in each calibration campaign.

Event selection requires an event to be fully reconstructed in all three coordinates ( $X$ ,  $Y$  and  $Z$ ). We compare the relative efficiency of this requirement for  $2\nu\beta\beta$  from MC to the measured relative efficiency derived from the background-subtracted low-background energy spectrum. Here, we define the relative efficiency as the ratio of the number of events passing the entire set of selection requirements to the number passing the set not including the full-reconstruction requirement. The relative efficiency from simulation changes modestly across the  $2\nu\beta\beta$  energy range (>99% to 90% from 980 keV to 2,450 keV) and similar behaviour is seen in data. The average deviation between simulation and data over the  $2\nu\beta\beta$  spectrum (7.8%) is taken as a systematic error on the efficiency.



**Figure 3 | Event multiplicity in data and simulation.** **a**, Plot of  $\text{SS}/(\text{SS} + \text{MS})$  ratio versus energy in data for  $^{226}\text{Ra}$ ,  $^{60}\text{Co}$  and  $^{228}\text{Th}$  calibration sources. **b**, Comparison of  $\text{SS}/(\text{SS} + \text{MS})$  ratio between data and simulation for the three sources, as a function of energy. Despite having different underlying energy spectra, all sources exhibit similar behaviour across the shown energy range when comparing data and simulation (**b**). Error bars,  $\pm 1$  s.d.

**Table 1 |  $0\nu\beta\beta$  signal efficiency and associated systematic errors**

Source	Signal efficiency (%)	Error (%)
Summary from ref. 9	93.1	0.9
Partial reconstruction	90.9	7.8
Fiducial volume/rate agreement	NA	3.4
Total	84.6	8.6

\*Partial reconstruction\* refers to the requirement that all events be fully reconstructed in  $X$ ,  $Y$  and  $Z$ . The summary for event selection from ref. 9 includes all efficiencies and related errors except fiducial volume and partial reconstruction, which have been recalculated in this work for  $0\nu\beta\beta$ . NA, not applicable.

The uncertainty on the location of the ROI in the spectrum is dominated by a possible energy-scale difference between  $\beta$ -like events in the LXe (for example,  $0\nu\beta\beta$ ) and  $\gamma$ -like events (including most backgrounds and the sources used for the primary energy calibration). We define the ‘ $\beta$ -scale’ as  $E_\beta = BE_\gamma$ , where  $E_\beta$  ( $E_\gamma$ ) is the energy for depositions from  $\beta$ s ( $\gamma$ s) and  $B$  is a measured constant. We determine the  $\beta$ -scale by fitting to the  $2\nu\beta\beta$ -decay-dominated low-background data and find  $B = 0.999 \pm 0.002$ .

Several cross-checks were performed to search for energy dependence in the  $\beta$ -scale. The above fits were performed using different energy thresholds and with different background PDFs produced using the skewing functions discussed earlier. We also fitted the low-background data assuming a linear energy dependence (for example,  $p_0 + p_1 E_\gamma$ ) for  $B$ . In all cases the results are consistent with the original fit, providing no evidence for energy dependence of the  $\beta$ -scale. The estimate of the  $\beta$ -scale is also robust against a different choice of  $2\nu\beta\beta$  spectral shape<sup>21</sup>.

To investigate the dependence of the ROI background estimate on the completeness of the model used to fit the data, we derive PDFs from different source locations and introduce them separately into the default background model used in the fit. The relative change of the estimated ROI background is then determined. The three background PDFs considered in this study are  $^{238}\text{U}$  in the HFE and inner cryostat, and  $^{60}\text{Co}$  in the copper source guide tube. These were chosen because the initial source location affects relative amplitudes and spectral features in the ROI, that is, the  $^{214}\text{Bi}$   $\gamma$  (2,448 keV) and  $^{60}\text{Co}$  sum peak. This study indicates a total possible deviation of 5.7% for the expected background counts in the ROI.

The residual time dependence of the energy resolution (Fig. 1) can introduce additional counts in the ROI from the 2,615-keV  $^{208}\text{Tl}$  peak. This was estimated to affect the ROI background counts by  $\pm 1.5\%$ .

A summary of the  $0\nu\beta\beta$  signal efficiency and associated uncertainty is presented in Table 1. Table 2 summarizes the uncertainties on the estimation of background in the ROI. These errors are explicitly included as input to the final fit to the low-background data. Items not listed in the tables, such as the  $\beta$ -scale and the SS/MS ratios, still contribute to the total systematic error on the  $0\nu\beta\beta$  signal as they are propagated to the final result by the maximum-likelihood fit to the low-background data.

Neutrons arising from cosmic-ray muons or radioactive decays in the salt surrounding the laboratory may contribute background to the  $0\nu\beta\beta$  ROI via neutron capture or spallation processes. The contribution in the ROI is expected to arise primarily from neutron-capture  $\gamma$ s in the LXe and surrounding materials (for example, capture on  $^{63}\text{Cu}$  and  $^{65}\text{Cu}$  in the copper components, and on  $^{136}\text{Xe}$  in the LXe). A simulation using a simplified experimental geometry and employing the FLUKA<sup>22,23</sup> and SOURCES<sup>24</sup> software packages is used to generate,

**Table 2 | Systematic errors on background determination in the ROI**

Source	Error (%)
Background shape distortion	9.2
Background model	5.7
Energy resolution variation	1.5
Total	10.9

These errors arise from incorrect modelling of the background shape (‘Background shape distortion’), incorrect or incomplete background model (‘Background model’) and the residual variation of the energy resolution over time (‘Energy resolution variation’; see, for example, Fig. 1).



track and thermalize neutrons. The resulting neutron capture rates are used as input to the Geant4-based<sup>20</sup> EXO-200 simulation package<sup>9</sup>, with the respective  $n$ -capture  $\gamma$ -spectra produced on the basis of ENSDF information<sup>25</sup> for the given nuclides. The produced PDFs are used in fits to the low-background data. Good shape agreement is found between these PDFs and data coincident with muon-veto-panel events.

## Results

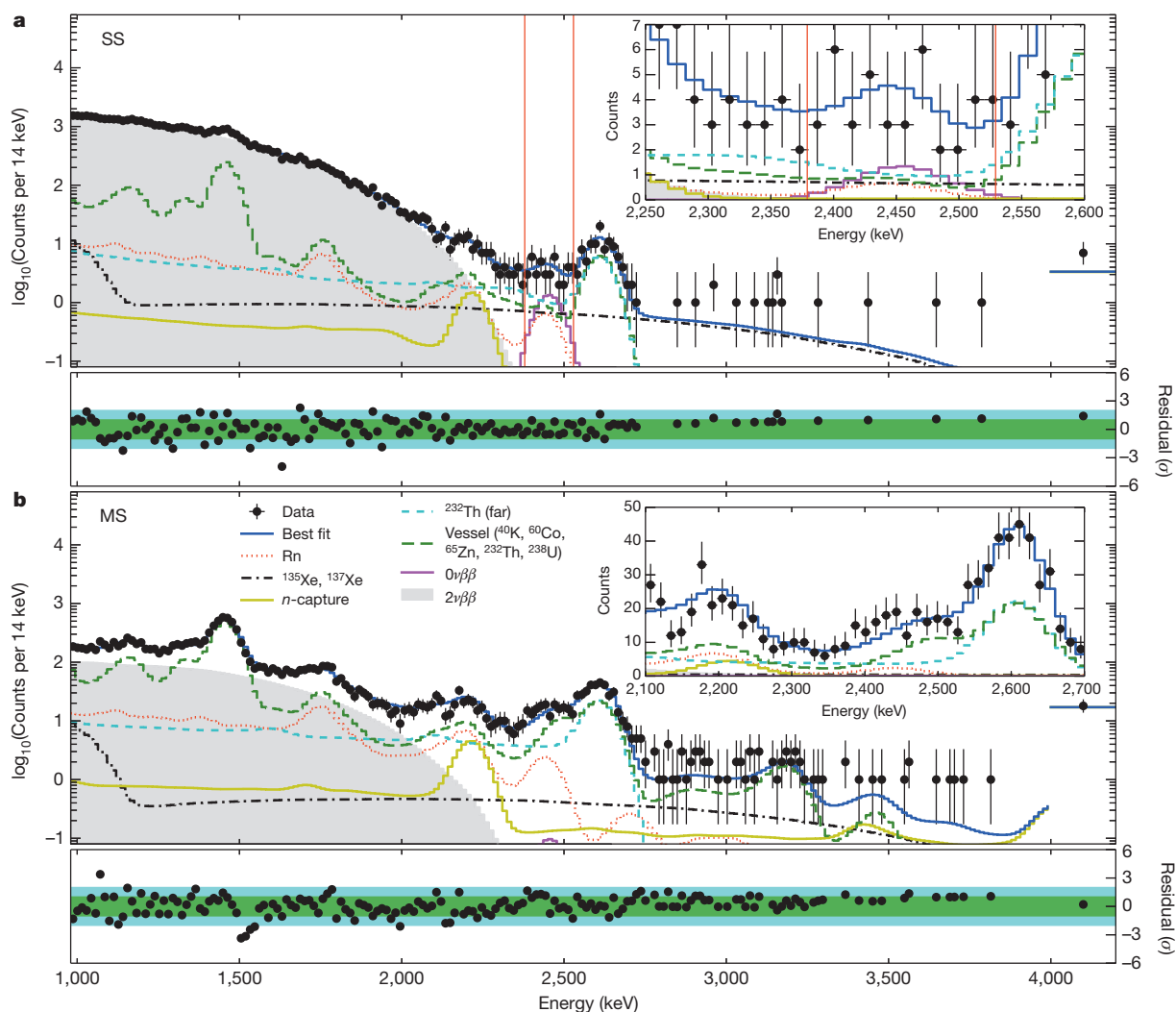
The fit to the low-background data minimizes the negative log-likelihood function constructed using a signal and background model composed of PDFs from simulation. A profile-likelihood scan is performed to search for a  $0\nu\beta\beta$  signal.

The PDFs chosen for the low-background fit model are those used in ref. 9 plus a ‘far-source’  $^{232}\text{Th}$  PDF, a  $^{137}\text{Xe}$  PDF and neutron-capture-related PDFs, including  $^{136}\text{Xe}$  neutron capture in the LXe,  $^1\text{H}$  neutron-capture in the HFE, and  $^{63}\text{Cu}$ ,  $^{65}\text{Cu}$  neutron capture in Cu components (LXe vessel, inner and outer cryostats). The far-source  $^{232}\text{Th}$  PDF allows for background contributions from Th in materials far from the TPC, for example in the HFE and in the copper cryostat. (Remote  $^{238}\text{U}$  is included in the fit model via  $^{222}\text{Rn}$ , simulated in the air between the

cryostat and Pb shield.) We combine the neutron-capture-related PDFs to form one PDF, allowing the relative rates of the component PDFs to float within 20% of their simulation-estimated values. The total rate of this summed PDF is allowed to float unconstrained.

We constrain the single-site fractions,  $\text{SS}/(\text{SS} + \text{MS})$ , of all components to be within 9.6% of their value calculated from simulation. An additional 90% correlation between single-site fractions of  $\gamma$  components is introduced into the likelihood function, owing to the consistent behaviour observed in these parameters in calibration studies (for example, Fig. 3). The overall normalization is allowed to float within the estimated systematic errors (8.6%). The background-PDF amplitudes within the ROI are also allowed to vary within their estimated systematic errors (10.9%). The  $\beta$ -scale is not allowed to float during the fit, but is manually profiled while performing the profile-likelihood scan for  $0\nu\beta\beta$ .

The final step before performing the fit was the unmasking of live-time around the SS ROI. However, before unmasking the full data set, we investigated backgrounds associated with Xe feeds, irregular occurrences in which additional Xe gas is introduced into the purification circulation loop. (These Xe feeds occurred 10 times over the run period



**Figure 4 | Fit results projected in energy.** **a, b,** Main panels show SS (**a**) and MS (**b**) events, as counts versus energy, with a zoom-in (inset) around the ROI: 2250–2600 keV (2100–2700 keV) for SS (MS); the bin size is 14 keV, and data points are shown in black. Lower panels in **a** and **b** show residuals between data and best fit normalized to the Poisson error, ignoring bins with 0 events. The green (blue) shaded regions in the lower panels represent  $\pm 1\sigma$  ( $\pm 2\sigma$ ) deviations. The 7 (18) events between 4,000 and 9,800 keV in the SS (MS)

spectrum have been collected into an overflow bin for presentation here. The vertical (red) lines in the SS spectra indicate the  $\pm 2\sigma$  ROI. The result of the simultaneous fit to the standoff distance is not shown here. Several background model components (including Rn,  $^{135}\text{Xe}$  and  $^{137}\text{Xe}$ ,  $n$ -capture,  $^{232}\text{Th}$  (far), Vessel,  $0\nu\beta\beta$  and  $2\nu\beta\beta$ , all described further in the text) are indicated in the main panel of **b** to show their relative contributions to the spectra. Error bars on data points,  $\pm 1$  s.d.





Received 27 February; accepted 28 April 2014.

Published online 4 June 2014.

- Majorana, E. Theory of the symmetry of electrons and positrons. *Nuovo Cimento* **14**, 171–184 (1937).
- Racah, G. On the symmetry of particles and antiparticles. *Nuovo Cimento* **14**, 322–328 (1937).
- Mourik, V. *et al.* Signatures of majorana fermions in hybrid superconductor-semiconductor nanowire devices. *Science* **336**, 1003–1007 (2012).
- Camilleri, L., Lisi, E. & Wilkerson, J. F. Neutrino masses and mixings: status and prospects. *Annu. Rev. Nucl. Part. Sci.* **58**, 343–369 (2008).
- Mohapatra, R. N. & Senjanovic, G. Neutrino mass and spontaneous parity violation. *Phys. Rev. Lett.* **44**, 912–915 (1980).
- Gell-Mann, M., Ramond, P. & Slansky, R. in *Supergravity* (eds van Nieuwenhuizen, P. & Freedman, D.) 315–321 (North Holland, 1979).
- Schechter, J. & Valle, J. Neutrinoless double- $\beta$  decay in SU(2)  $\times$  U(1) theories. *Phys. Rev. D* **25**, 2951–2954 (1982).
- Barabash, A. Precise half-life values for two-neutrino double- $\beta$  decay. *Phys. Rev. C* **81**, 035501 (2010).
- Albert, J. B. *et al.* (EXO-200 Collaboration). An improved measurement of the  $2\nu\beta\beta$  half-life of Xe-136 with EXO-200. *Phys. Rev. C* **89**, 015502 (2014).
- Vogel, P. Nuclear structure and double beta decay. *J. Phys. G* **39**, 124002 (2012).
- Agostini, M. *et al.* (GERDA Collaboration). Results on neutrinoless double- $\beta$  decay of  $^{76}\text{Ge}$  from Phase I of the GERDA experiment. *Phys. Rev. Lett.* **111**, 122503 (2013).
- Gando, A. *et al.* (KamLAND-Zen Collaboration). Limit on neutrinoless  $\beta\beta$  decay of  $^{136}\text{Xe}$  from the first phase of KamLAND-Zen and comparison with the positive claim in  $^{76}\text{Ge}$ . *Phys. Rev. Lett.* **110**, 062502 (2013).
- Auger, M. *et al.* (EXO-200 Collaboration). Search for neutrinoless double-beta decay in  $^{136}\text{Xe}$  with EXO-200. *Phys. Rev. Lett.* **109**, 032505 (2012).
- Klapdor-Kleingrothaus, H. & Krivosheina, I. The evidence for the observation of  $0\nu\beta\beta$  decay: the identification of  $0\nu\beta\beta$  events from the full spectra. *Mod. Phys. Lett. A* **21**, 1547–1566 (2006).
- Auger, M. *et al.* The EXO-200 detector, part I: Detector design and construction. *J. Instrum.* **7**, P05010 (2012).
- Redshaw, M., Wingfield, E., McDaniel, J. & Myers, E. G. Mass and double-beta-decay  $Q$  value of  $^{136}\text{Xe}$ . *Phys. Rev. Lett.* **98**, 062503 (2007).
- 3M HFE-7000 <http://www.3m.com> (2014).
- Esch, E.-I. The cosmic ray muon flux at WIPP. *Nucl. Instrum. Meth. A* **538**, 516–525 (2005).
- Leonard, D. *et al.* Systematic study of trace radioactive impurities in candidate construction materials for EXO-200. *Nucl. Instrum. Meth. A* **591**, 490–509 (2008).
- Allison, J. *et al.* Geant4 developments and applications. *IEEE Trans. Nucl. Sci.* **53**, 270–278 (2006).
- Kotila, J. & Iachello, F. Phase-space factors for double- $\beta$  decay. *Phys. Rev. C* **85**, 034316 (2012).
- Battistoni, G. *et al.* in *Hadronic Shower Simulation Workshop* (eds Albrow, M. & Raja, R.) 31–49 (American Institute of Physics Conf. Ser. Vol. 896, (2007).
- Ferrari, A., Sala, P. R., Fasso, A. & Ranft, J. *FLUKA: A Multi-Particle Transport Code (Program version 2005)* (Tech. Rep. CERN-2005-010, SLAC-R-773, INFN-TC-05-11, CERN, 2005).
- Wilson, W. B. *SOURCES 4A: A Code for Calculating ( $\alpha, n$ ) Spontaneous Fission, and Delayed Neutron Sources and Spectra* (Tech. Rep. LA-13639-MS, Los Alamos National Laboratory, 1999).
- Bhat, M. in *Nuclear Data for Science and Technology* (ed. Qaim, S.) 817–821 (Research Reports in Physics, Springer, 1992).
- Rodríguez, T. R. & Martínez-Pinedo, G. Energy density functional study of nuclear matrix elements for neutrinoless  $\beta\beta$  decay. *Phys. Rev. Lett.* **105**, 252503 (2010).
- Menéndez, J., Poves, A., Caurier, E. & Nowacki, F. Disassembling the nuclear matrix elements of the neutrinoless  $\beta\beta$  decay. *Nucl. Phys. A* **818**, 139–151 (2009).
- Barea, J., Kotila, J. & Iachello, F. Nuclear matrix elements for double- $\beta$  decay. *Phys. Rev. C* **87**, 014315 (2013).
- Šimkovic, F., Rodin, V., Faessler, A. & Vogel, P.  $0\nu\beta\beta$  and  $2\nu\beta\beta$  nuclear matrix elements, quasiparticle random-phase approximation, and isospin symmetry restoration. *Phys. Rev. C* **87**, 045501 (2013).
- Forero, D. V., Tórtola, M. & Valle, J. W. F. Global status of neutrino oscillation parameters after Neutrino-2012. *Phys. Rev. D* **86**, 073012 (2012).
- Ackerman, N. *et al.* (EXO-200 Collaboration). Observation of two-neutrino double-beta decay in  $^{136}\text{Xe}$  with EXO-200. *Phys. Rev. Lett.* **107**, 212501 (2011).

- Wilks, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* **9**, 60–62 (1938).
- Cowan, G. *Statistical Data Analysis* (Oxford Science Publications, Clarendon, 1998).

**Acknowledgements** EXO-200 is supported by the DOE and NSF in the United States, NSERC in Canada, SNF in Switzerland, NRF in Korea, RFBR (12-02-12145) in Russia and the DFG Cluster of Excellence ‘Universe’ in Germany. EXO-200 data analysis and simulation used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the US DOE under contract no. DE-AC02-05CH11231. The EXO-200 collaboration acknowledges the WIPP for their hospitality.

**Author Contributions** Each of the authors participated in the collection and analysis of the data reported here, with the following exceptions: D.B. contributed to the slow controls system; G.F.C. performed energy resolution simulations; X.S.J. and Y.B.Z. provided electronics expertise; M. Danilov, A.D., T.K. and P.V. contributed to the initial conception and design of the experiment; M. Danilov and A.D. also contributed to the acquisition of the xenon, while P.V. also advised on nuclear and particle theory; J.D., R.N. and A.R. provided engineering, operations and technical support at the WIPP facility; A.J., J.J.R. and A.W. supported data acquisition, data processing and software. In line with collaboration policy, the authors are listed here alphabetically. EXO-200 was constructed and commissioned by the authors of refs 15 and 31.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G.M. (michael.marino@mytum.de).

### The EXO-200 Collaboration

J. B. Albert<sup>1</sup>, D. J. Auty<sup>2</sup>, P. S. Barbeau<sup>3</sup>, E. Beauchamp<sup>4</sup>, D. Beck<sup>5</sup>, V. Belov<sup>6</sup>, C. Benitez-Medina<sup>7</sup>, J. Bonatt<sup>8,9</sup>, M. Breidenbach<sup>10</sup>, T. Brunner<sup>9</sup>, A. Burenkov<sup>6</sup>, G. F. Cao<sup>11</sup>, C. Chambers<sup>7</sup>, J. Chaves<sup>9</sup>, B. Cleveland<sup>4,12</sup>, M. Coon<sup>5</sup>, A. Craycraft<sup>7</sup>, T. Daniels<sup>8</sup>, M. Danilov<sup>6</sup>, S. J. Daugherty<sup>1</sup>, C. G. Davis<sup>13</sup>, J. Davis<sup>9</sup>, R. DeVoe<sup>9</sup>, S. Delaquis<sup>14</sup>, T. Didberidze<sup>2</sup>, A. Dolgolenko<sup>6</sup>, M. J. Dolinski<sup>15</sup>, M. Dunford<sup>16</sup>, W. Fairbank Jr<sup>7</sup>, J. Farine<sup>4</sup>, W. Feldmeier<sup>17</sup>, P. Fierlinger<sup>17</sup>, D. Fudenberg<sup>9</sup>, G. Giroux<sup>14</sup>, R. Gornea<sup>14</sup>, K. Graham<sup>16</sup>, G. Gratta<sup>9</sup>, C. Hall<sup>13</sup>, S. Herrin<sup>10</sup>, M. Hughes<sup>2</sup>, M. J. Jewell<sup>15</sup>, X. S. Jiang<sup>11</sup>, A. Johnson<sup>10</sup>, T. N. Johnson<sup>1</sup>, S. Johnston<sup>8</sup>, A. Karelin<sup>6</sup>, L. J. Kaufman<sup>1</sup>, R. Killick<sup>16</sup>, T. Koffas<sup>16</sup>, S. Kravitz<sup>9</sup>, A. Kuchenkov<sup>6</sup>, K. S. Kumar<sup>8</sup>, D. S. Leonard<sup>18</sup>, F. Leonard<sup>16</sup>, C. Licciardi<sup>16</sup>, Y. H. Lin<sup>15</sup>, R. MacLellan<sup>10</sup>, M. G. Marino<sup>17</sup>, B. Mong<sup>4</sup>, D. Moore<sup>9</sup>, R. Nelson<sup>19</sup>, A. Odian<sup>10</sup>, I. Ostrovskiy<sup>9</sup>, C. Ouellet<sup>16</sup>, A. Piepke<sup>2</sup>, A. Pocar<sup>8</sup>, C. Y. Prescott<sup>10</sup>, A. Rivas<sup>9</sup>, P. C. Rowson<sup>10</sup>, M. P. Roze<sup>16</sup>, J. J. Russell<sup>10</sup>, A. Schubert<sup>9</sup>, D. Sinclair<sup>16,20</sup>, S. Slutsky<sup>13</sup>, E. Smith<sup>15</sup>, V. Stekhanov<sup>6</sup>, M. Tarka<sup>5</sup>, T. Tolba<sup>14</sup>, D. Tosi<sup>9</sup>, K. Twelker<sup>9</sup>, P. Vogel<sup>21</sup>, J.-L. Vuilleumier<sup>14</sup>, A. Waite<sup>10</sup>, J. Walton<sup>5</sup>, T. Walton<sup>7</sup>, M. Weber<sup>9</sup>, L. J. Wen<sup>11</sup>, U. Wichoski<sup>4</sup>, J. D. Wright<sup>8</sup>, L. Yang<sup>2</sup>, Y.-R. Yen<sup>13,15</sup>, O. Ya. Zeldovich<sup>6</sup> & Y. B. Zhao<sup>11</sup>

<sup>1</sup>Physics Department and CEEM, Indiana University, Bloomington, Indiana 47405, USA. <sup>2</sup>Department of Physics and Astronomy, University of Alabama, Tuscaloosa, Alabama 35487, USA. <sup>3</sup>Department of Physics, Duke University, and Triangle Universities Nuclear Laboratory (TUNL), Durham, North Carolina 27708, USA. <sup>4</sup>Department of Physics, Laurentian University, Sudbury, Ontario P3E 2C6, Canada. <sup>5</sup>Physics Department, University of Illinois, Urbana-Champaign, Illinois 61801, USA. <sup>6</sup>Institute for Theoretical and Experimental Physics, Moscow 117218, Russia. <sup>7</sup>Physics Department, Colorado State University, Fort Collins, Colorado 80523, USA. <sup>8</sup>Physics Department, University of Massachusetts, Amherst, Massachusetts 01003, USA. <sup>9</sup>Physics Department, Stanford University, Stanford, California 94305, USA. <sup>10</sup>SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. <sup>11</sup>Institute of High Energy Physics, Beijing 100049, China. <sup>12</sup>SNOLAB, Sudbury, Ontario P3Y 1N2, Canada. <sup>13</sup>Physics Department, University of Maryland, College Park, Maryland 20742, USA. <sup>14</sup>LHEP, Albert Einstein Center, University of Bern, 3012 Bern, Switzerland. <sup>15</sup>Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, USA. <sup>16</sup>Physics Department, Carleton University, Ottawa, Ontario K1S 5B6, Canada. <sup>17</sup>Technische Universität München, Physikdepartment, and Excellence Cluster Universe, 85748 Garching, Germany. <sup>18</sup>Department of Physics, University of Seoul, Seoul 130-743, South Korea. <sup>19</sup>Waste Isolation Pilot Plant, Carlsbad, New Mexico 88220, USA. <sup>20</sup>TRIUMF, Vancouver, British Columbia V6T 2A3, Canada. <sup>21</sup>Kellogg Laboratory, Caltech, Pasadena, California 91125, USA. <sup>†</sup>Queen's University, Kingston, Ontario K7L 3N6, Canada (G. Giroux); California Institute of Technology, Pasadena, California 91125, USA (S.S.); University of Wisconsin, Madison, Wisconsin 53706, USA (D.T.).

# Targeted genome editing in human repopulating haematopoietic stem cells

Pietro Genovese<sup>1</sup>, Giulia Schirotti<sup>1,2</sup>, Giulia Escobar<sup>1,2</sup>, Tiziano Di Tomaso<sup>1</sup>, Claudia Firrito<sup>1</sup>, Andrea Calabria<sup>1</sup>, Davide Moi<sup>1,†</sup>, Roberta Mazzieri<sup>1,†</sup>, Chiara Bonini<sup>3</sup>, Michael C. Holmes<sup>4</sup>, Philip D. Gregory<sup>4</sup>, Mirjam van der Burg<sup>5</sup>, Bernhard Gentner<sup>1,2</sup>, Eugenio Montini<sup>1</sup>, Angelo Lombardo<sup>1,2,\*</sup> & Luigi Naldini<sup>1,2,\*</sup>

**Targeted genome editing by artificial nucleases has brought the goal of site-specific transgene integration and gene correction within the reach of gene therapy. However, its application to long-term repopulating haematopoietic stem cells (HSCs) has remained elusive. Here we show that poor permissiveness to gene transfer and limited proficiency of the homology-directed DNA repair pathway constrain gene targeting in human HSCs. By tailoring delivery platforms and culture conditions we overcame these barriers and provide stringent evidence of targeted integration in human HSCs by long-term multilineage repopulation of transplanted mice. We demonstrate the therapeutic potential of our strategy by targeting a corrective complementary DNA into the *IL2RG* gene of HSCs from healthy donors and a subject with X-linked severe combined immunodeficiency (SCID-X1). Gene-edited HSCs sustained normal haematopoiesis and gave rise to functional lymphoid cells that possess a selective growth advantage over those carrying disruptive *IL2RG* mutations. These results open up new avenues for treating SCID-X1 and other diseases.**

Haematopoietic stem cell (HSC)-based gene therapy has provided therapeutic benefit in primary immunodeficiencies<sup>1,2</sup>, thalassaemia<sup>3</sup> and leukodystrophies<sup>4,5</sup>. Whereas more advanced vectors, such as lentiviral vectors<sup>6</sup>, have shown improved safety and efficacy, the risk of insertional mutagenesis<sup>7,8</sup> and unregulated transgene expression<sup>9,10</sup> remains a concern when using semi-randomly integrating vectors. These adverse effects may trigger oncogenesis, toxicity or elimination of the gene-modified cells.

Artificial endonucleases, such as zinc finger nucleases (ZFNs)<sup>11</sup>, transcription activator-like effector nucleases (TALENs)<sup>12</sup>, and RNA-guided nucleases (CRISPR/Cas)<sup>13</sup>, brought the possibility of gene targeting within the reach of gene therapy<sup>14,15</sup>. These nucleases are used to efficiently and specifically target a DNA double-strand break (DSB) to a pre-selected genomic site<sup>13,16,17</sup>. According to the repair process that seals the break<sup>18</sup>, the outcome can be disruption, reconstitution or editing of the original sequence. If the DSB is sealed by the error-prone non-homologous end-joining (NHEJ) pathway, insertions and deletions (indels) are common<sup>19–21</sup>. If the DSB is sealed by the high-fidelity homology directed repair (HDR) pathway, which acts preferentially during the S/G2 phase, the targeted sequence can be edited by providing an exogenous DNA template flanked by homologous sequences to the nuclease target site. Targeted editing allows for the integration of an expression cassette into a safe genomic harbour<sup>22,23</sup>, or correcting disease-causing mutations by inserting a functional copy of the affected gene downstream of its own promoter<sup>14,24</sup>. Gene correction, as opposed to gene replacement, may not only restore the function but also the physiological expression of the gene, a long-sought goal of gene therapy.

Whereas gene disruption by ZFNs has been shown in human haematopoietic stem/progenitor cells (HSPCs) assayed by repopulation of SCID mice<sup>20</sup>, targeted gene editing in these cells has not been reported. Here we identify and overcome major constraints limiting gene targeting in HSPCs and provide proof-of-efficacy for this approach by functional reconstitution of the *IL2RG* gene, mutations of which are responsible for SCID-X1.

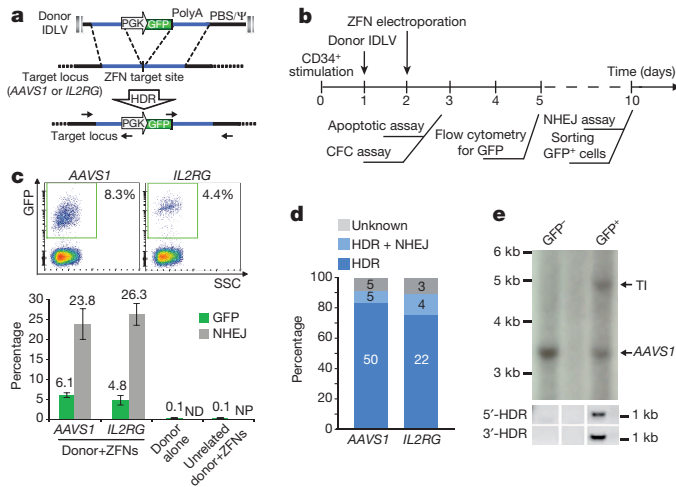
## Site-specific integration in human HSPCs

We developed a protocol for delivery of ZFNs and donor DNA template into human umbilical cord blood CD34<sup>+</sup> cells by mRNA electroporation and integrase-defective lentiviral vector (IDLV), and targeted integration of a GFP cassette into the *AAVS1* 'safe harbour'<sup>22</sup> or a mutational hotspot of *IL2RG* (refs 14, 15) (Fig. 1a, b and Extended Data Fig. 1). This protocol yielded on average 5% GFP<sup>+</sup> cells in liquid culture and colony-forming cell (CFC) assays and high percentages of indels in the respective ZFN target sites, although after a transient cell loss (Fig. 1c and Extended Data Fig. 2a–e). PCR analysis and Southern blot showed integration of the GFP cassette at the intended targets in >90% of the GFP<sup>+</sup> colonies ( $n = 89$ ) and in induced pluripotent stem cells obtained by reprogramming the GFP<sup>+</sup> sorted cells (Fig. 1d, e and Extended Data Fig. 2f, g). We then transplanted the CD34<sup>+</sup> cells treated for *AAVS1* or *IL2RG* gene targeting into NOD-SCID-*Il2rg*<sup>−/−</sup> (NSG) mice and found human cell engraftment in all mice (Fig. 2a). In the first 8 weeks after transplant, 95% of the mice had circulating GFP<sup>+</sup> cells (mean  $6.2 \pm 1.3\%$ ; Fig. 2b), whereas only 42% of the mice maintained long-term GFP marking. End-point analyses performed on the peripheral blood, spleen and bone marrow showed that GFP<sup>+</sup> cells were present within all human haematopoietic lineages, including lymphoid and myeloid cells, and erythroid precursors (mean  $2 \pm 0.8\%$ ; Fig. 2c). Similar frequencies of GFP<sup>+</sup> cells were found among primitive and committed progenitors in the bone marrow (mean  $2.2 \pm 0.9\%$ ; Fig. 2d). PCR analyses confirmed targeted integration in human lymphoid, myeloid and CD34<sup>+</sup> cells from the spleen and bone marrow of representative mice (Fig. 2e). CFC assays on CD34<sup>+</sup> cells from bone marrow showed GFP<sup>+</sup> myeloid and erythroid colonies (Fig. 2f) with targeted integration (Fig. 2e). Analysis of bone marrow cells showed the occurrence of NHEJ-mediated indels in the ZFN target site in the majority of mice (64%; Fig. 2g) at higher levels than observed for GFP marking, indicating that DNA DSB induction and repair by either HDR or NHEJ is compatible with haematopoietic repopulation. Overall, these data show that our gene targeting

<sup>1</sup>TIGET, San Raffaele Telethon Institute for Gene Therapy, San Raffaele Scientific Institute, 20132 Milan, Italy. <sup>2</sup>Vita Salute San Raffaele University, 20132 Milan, Italy. <sup>3</sup>Experimental Hematology Unit, San Raffaele Scientific Institute, 20132 Milan, Italy. <sup>4</sup>Sangamo BioSciences Inc., Richmond, California 94804, USA. <sup>5</sup>Department of Immunology Erasmus MC, University Medical Center, 3015 Rotterdam, The Netherlands. <sup>†</sup>Present address: The University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland 4102, Australia.

\*These authors contributed equally to this work.





**Figure 1 | Targeted integration into *AAVS1* or *IL2RG* in umbilical cord blood  $CD34^+$  cells.** **a**, Schematic of the donor IDLV template containing a GFP cassette driven by the phosphoglycerate kinase promoter (PGK) flanked by sequences homologous to the genomic target locus, the target locus with the ZFN cleavage site and the locus after HDR showing the PCR primers used to assess 5' or 3' HDR-mediated integration junctions (black arrows). **b**, Flow chart for targeted integration and cell analyses. **c**, Representative flow cytometry dot plots (top) and percentages of GFP<sup>+</sup> cells and NHEJ-induced indels at the target locus (bottom) of cord blood  $CD34^+$  cells treated for targeted integration into *AAVS1* or *IL2RG*. Mean  $\pm$  s.e.m. (*AAVS1*,  $n = 39$  on 19 cord blood donors; *IL2RG*,  $n = 10$  on 9 cord blood donors). Unrelated donor indicates cells treated with IDLV lacking homology to the target site. ND, not detectable; NP, not performed. **d**, Targeting specificity in CFCs. Percentage of colonies positive for both (HDR), either (HDR+NHEJ) or none (Unknown) 5' and 3' HDR junctions by PCR. Numbers of colonies screened are indicated inside the bars. **e**, Southern blot (top) and PCR (bottom) analyses of iPSCs obtained by reprogramming GFP<sup>+</sup> or GFP<sup>-</sup> cells from panel **c**. TI, targeted integration.

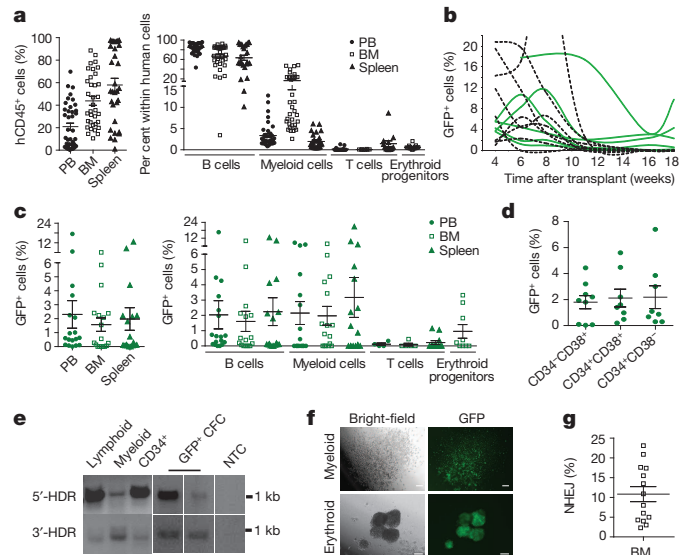
protocol achieves site-specific integration in human multipotent long-term NSG repopulating cells (SRCs), surrogate readouts of HSCs.

### Low targeting efficiency in HSPCs

The *in vivo* studies revealed that only ~40% of mice had on average 2% human GFP<sup>+</sup> cells over the long term. These figures seem lower than expected from transplanting  $CD34^+$  cells with ~5% targeting efficiency *in vitro* (Supplementary Information) and suggest that either SRCs are targeted less efficiently than the bulk  $CD34^+$  cells, or the gene-targeted SRCs have a competitive disadvantage *in vivo*. We thus compared the percentages of GFP<sup>+</sup> cells among different subpopulations of cultured  $CD34^+$  cells, prospectively identified by surface markers<sup>23</sup> as primitive ( $CD34^+CD133^+CD90^+$ ), early ( $CD34^+CD133^+CD90^-$ ) and committed ( $CD34^+CD133^-$ ) progenitors, and the differentiated cells ( $CD34^-$ , Fig. 3a, left panel). We found a decreasing frequency of GFP<sup>+</sup> cells when moving from the differentiated cells up the progenitor hierarchy. In primitive cells the percentage of GFP<sup>+</sup> cells was 20-fold lower than that measured in differentiated cells. We thus investigated the potential rate-limiting steps for gene targeting in primitive cells (Extended Data Fig. 3a–d). Transgene expression on mRNA electroporation was similar among the subpopulations or slightly lower for the primitive cells. The level of NHEJ induced at the ZFN target site was higher in the primitive cells and progressively lower in committed and differentiated cells. This difference, however, diminished with time in culture, potentially due to the loss of some treated primitive cells. Indeed, induction of apoptosis was higher in this subset. Taken together these data indicate that the primitive cells are more sensitive to our treatment and less permissive to HDR and/or donor template delivery.

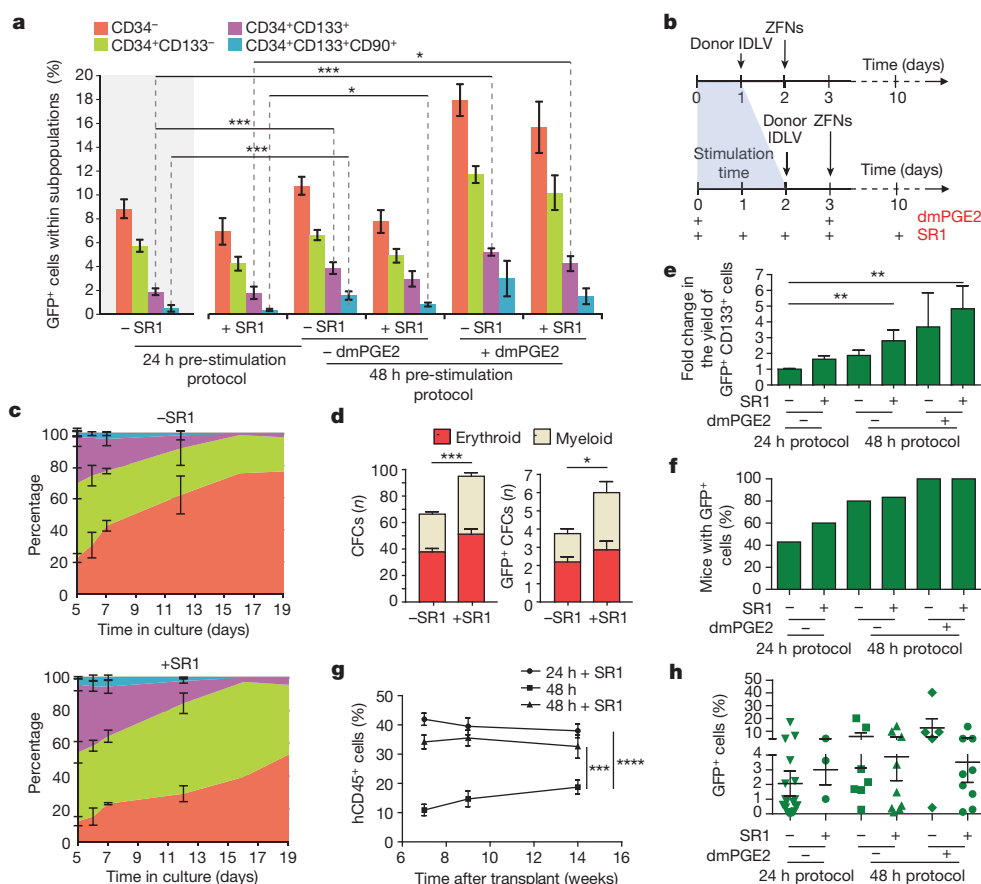
### Tailored conditions improve HSC gene targeting

Because cell-cycle progression is a requirement of HDR and activation of the primitive progenitors may require longer stimulation, we postponed



**Figure 2 | Transplantation of gene-targeted  $CD34^+$  cells in NSG mice.**  $CD34^+$  cells treated as in Fig. 1b were transplanted into NSG mice. **a**, Left: human cell engraftment ( $CD45^+$ ) 12–23 weeks after transplant in the indicated organs. BM, bone marrow; PB, peripheral blood. Right: percentage of the indicated lineages within the human graft. Data from individual mice and mean  $\pm$  s.e.m. ( $n = 42$  mice; 6 independent experiments on 13 cord blood donors). **b**, Time course of human GFP<sup>+</sup> cells in peripheral blood (PB) of mice. Dashed lines indicate mice in which GFP<sup>+</sup> cells were no longer detectable (<0.1%) 12 weeks after transplant. **c**, GFP<sup>+</sup> cells within the human graft in the indicated organs (left) and lineages (right) ( $n = 18$ ). **d**, GFP<sup>+</sup> cells within human primitive ( $CD34^+CD38^-$ ) or committed ( $CD34^+CD38^+$ ) progenitors or differentiated cells ( $CD34^-CD38^+$ ) in mouse bone marrow. **e**, PCR analysis for targeted integration into *AAVS1* on human lymphoid ( $CD19^+$ ) and myeloid ( $CD33^+$  and  $CD13^+$ ) cells sorted from the mice and on GFP<sup>+</sup> CFCs from mouse bone marrow. **f**, Representative images of GFP<sup>+</sup> colonies. Scale bars, 0.5 mm. **g**, NHEJ at *AAVS1* or *IL2RG* ZFN target sites on total bone marrow cells from panel **a**.  $n = 25$ , 3 independent experiments.

the gene targeting procedure to the third day of culture (Fig. 3b). At this time, the cells are also likely to become more permissive to lentiviral vector transduction. Because increasing the length of time in culture promotes differentiation, we added the aryl hydrocarbon receptor antagonist (StemRegenin 1, SR1)<sup>26</sup> and/or 16,16-dimethyl-prostaglandin E2 (dmPGE2)<sup>27,28</sup> to the culture to better preserve stem and early progenitor cells (Fig. 3c). The delayed protocol significantly increased ( $\geq 2$ -fold) the percentage of GFP<sup>+</sup> cells in primitive cells (Fig. 3a, right panels). SR1 slightly reduced the percentage of GFP<sup>+</sup> cells observed within each subpopulation but increased the yield of GFP<sup>+</sup> CFCs and early progenitors consistently with the increased proportion of immature cells in SR1-treated cultures (Fig. 3d, e and Extended Data Fig. 3e). The addition of dmPGE2 increased the percentage of GFP<sup>+</sup> cells in all subpopulations when used alone and showed additive effects with SR1. Notably, both the delayed treatment and the addition of SR1 and dmPGE2 increased the fraction of mice showing long-term engraftment with GFP<sup>+</sup> cells, which reached 100% when used in combination (Fig. 3f). Human cell engraftment significantly increased after addition of SR1 to the culture (Fig. 3g) and was stable long-term (Extended Data Fig. 4a, b). Consistent with the increased GFP marking observed *in vitro* in the primitive cells, the mean percentage of GFP<sup>+</sup> cells showing long-term engraftment *in vivo* increased with all types of delayed treatments (Fig. 3h). GFP<sup>+</sup> cells contributed to multiple lineages and to the progenitor compartment in most mice (Extended Data Fig. 4c, d). Molecular analyses on bone marrow cells showed evidence of targeted integration (Extended Data Fig. 4e). Serial transplant of purified  $CD34^+$  cells from the bone marrow of primary mice showed engraftment and differentiation of targeted GFP<sup>+</sup> cells in secondary recipients (Extended Data Fig. 4f, g).



**Figure 3 | Gene targeting in primitive versus committed progenitors.** **a**, GFP<sup>+</sup> cells within the indicated subpopulations 3 days after treatment for targeted integration. The left-most panel shows results using the protocol described in Fig. 1b. The other panels show the effect of longer pre-stimulation and/or addition of the indicated drugs, as shown in the schematic in panel **b**. Means  $\pm$  s.e.m. ( $n = 31, 15, 14, 15, 7, 5$  respectively on 37 total cord blood donors). \* $P < 0.05$ ; \*\*\* $P < 0.001$  (one-way ANOVA). **c**, Composition of CD34<sup>+</sup> cells cultured with or without SR1; subpopulations as in **a**. Means  $\pm$  s.e.m. ( $n = 4$ ). **d**, Total (left) and GFP<sup>+</sup> (right) colonies from CD34<sup>+</sup> cells treated for targeted integration with or without SR1. Means  $\pm$  s.e.m.

( $n = 20, 14$ ). **e**, Yield of GFP<sup>+</sup> early progenitors relative to that obtained using the original protocol of Fig. 1b. Means  $\pm$  s.e.m. ( $n = 8, 7, 11, 10, 3, 5$ ) \*\* $P < 0.01$  (one-way ANOVA). **f**, Percentage of NSG mice harbouring GFP<sup>+</sup> cells 14 weeks after transplant of CD34<sup>+</sup> cells treated with the indicated targeted integration protocols. **g**, Time course of human engraftment in peripheral blood. Means  $\pm$  s.e.m. (24 h SR1,  $n = 5$ ; 48 h SR1,  $n = 6$ ; 48 h,  $n = 5$ ) \*\*\*\* $P < 0.0001$ , \*\*\* $P < 0.001$  (two-way ANOVA). **h**, GFP<sup>+</sup> cells within CD45<sup>+</sup> cells in peripheral blood 14 weeks after transplant. Means  $\pm$  s.e.m. ( $n = 4$ ). Mice for the 24 h SR1<sup>-</sup> condition are shown for comparison from Fig. 2c.

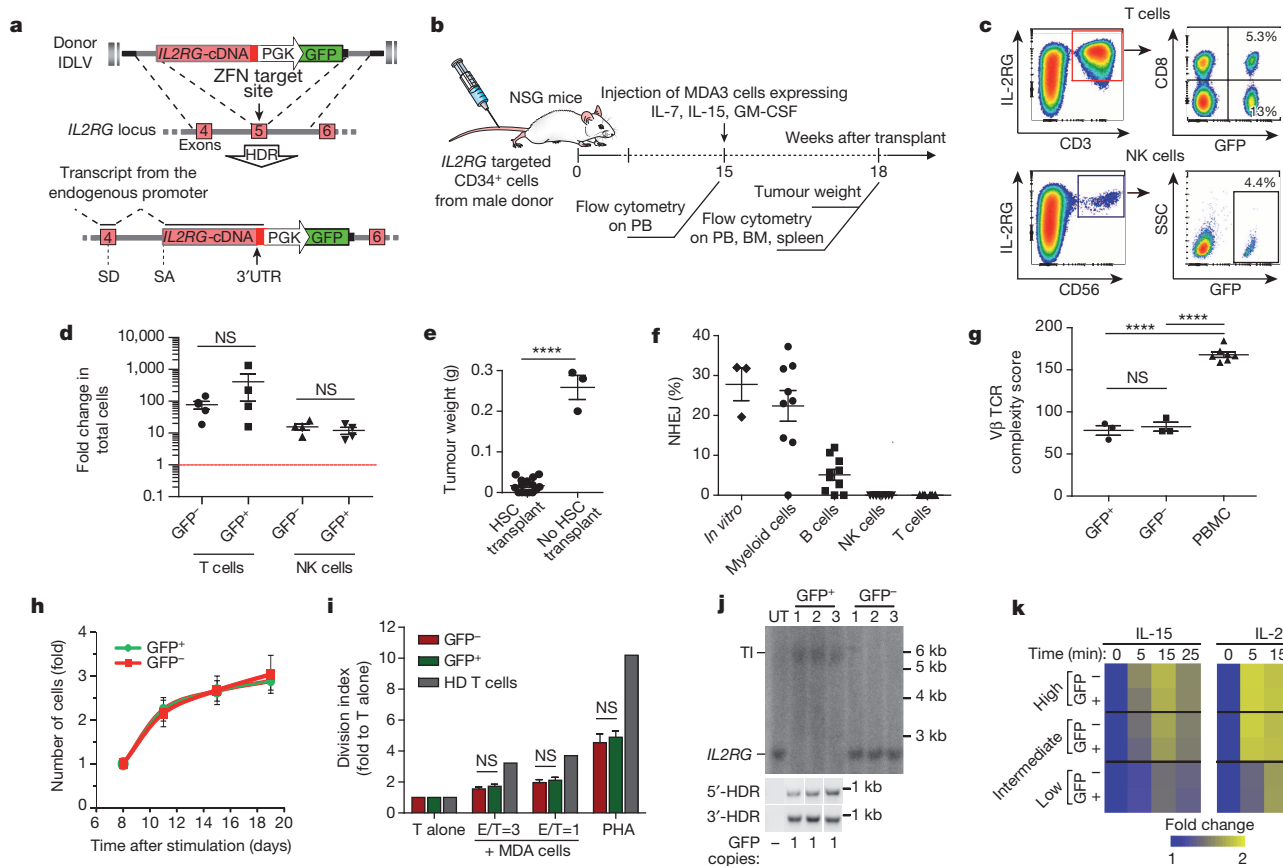
Overall, these data indicate that by tailoring experimental conditions we could improve the yield and frequency of targeted long-term SRCs.

### Targeted gene editing of *IL2RG* in HSPCs

In the experiments described in Fig. 3g, h, the gene targeting construct was designed to insert a cDNA comprising exons 5–8 of *IL2RG* together with the GFP cassette into the *IL2RG* gene and used on CD34<sup>+</sup> cells from healthy male donors (Fig. 4a). In this way, the cDNA is transcribed from the endogenous *IL2RG* promoter and spliced to its upstream exons, thus providing a platform for correcting all SCID-X1-causing mutations downstream of exon 4. To assess functional reconstitution of the targeted gene, we challenged the repopulated mice with a human tumour cell line (MDA-MB 231) engineered to express human interleukin (IL)-7, IL-15 and granulocyte-macrophage colony-stimulating factor (GM-CSF) (Fig. 4b). We previously reported that this challenge leads to improved reconstitution of functional human T and natural killer (NK) cells that eventually reject the tumour graft<sup>29</sup>. T and NK cells are strictly dependent on IL-2RG expression for survival and activity and are absent in SCID-X1. On tumour challenge, we observed a massive (mean  $130 \pm 40$ -fold) expansion of the human T and NK lineages in the repopulated mice (Extended Data Fig. 5). GFP<sup>+</sup> T and NK cells expressed IL-2RG on the cell surface (Fig. 4c) and expanded similarly to their GFP-negative counterparts in all mice (Fig. 4d). All repopulated mice effectively rejected the allogeneic tumour, at variance with non-transplanted mice,

underscoring the development of a functional human immune system (Fig. 4e). Whereas myeloid cells sorted from the mice showed high levels of NHEJ at the targeted *IL2RG* site, comparable to those observed in the CD34<sup>+</sup> cells pre-transplant, B cells showed very little NHEJ, and T and NK cells showed virtually no NHEJ (Fig. 4f). These findings reflect the marked counter selection of lymphoid cells carrying a disrupted *IL2RG*, as it naturally occurs with inherited SCID-X1 alleles, and confirm the functionality of the reconstituted gene in the expanded GFP<sup>+</sup> cells. We then assessed the T-cell-receptor repertoire of lymphocytes from the engrafted mice and found substantial diversity with almost overlapping polyclonal pattern between the GFP<sup>+</sup> and GFP<sup>-</sup> sorted cell subsets (Fig. 4g and Extended Data Fig. 6). The GFP<sup>+</sup> and GFP<sup>-</sup> T cells expanded *ex vivo* after polyclonal stimulation with the same kinetics in the presence of  $\gamma$ -chain-dependent cytokines (IL-7 and IL-15), and proliferated to a similar extent in response to the allogeneic MDA-MB 231 cells (Fig. 4h, i and Extended Data Fig. 7a). GFP<sup>+</sup> and GFP<sup>-</sup> T cells were similarly comprised of CD8 and CD4 cell subsets, with most cells showing effector phenotypes (Extended Data Fig. 7b, c). Consistently, both GFP<sup>+</sup> and GFP<sup>-</sup> cells robustly produced IFN- $\gamma$  and IL-2 after phorbol myristate acetate (PMA)-ionomycin stimulation or when co-cultured with the allogeneic tumour at different effector/target ratios (Extended Data Fig. 7d, e). Molecular analyses demonstrated that nearly all GFP<sup>+</sup> cells contained targeted integration into *IL2RG* (Fig. 4j). We then measured the phosphorylation of two downstream effectors in the signalling





**Figure 4 | Functional reconstitution of *IL2RG* in the lymphoid progeny of HSCs.** **a**, Schematic of the *IL2RG* donor: a promoter-less *IL2RG* cDNA, comprising exons 5–8 plus 3' untranslated region (UTR), and a PGK-GFP cassette are flanked by homologous sequences to those surrounding the *IL2RG* ZFN target site. SA, splice acceptor; SD, splice donor. **b**, Flow chart of cell transplantation, tumour challenge and analyses. **c**, Density plots of  $\gamma$ -chain-expressing T (top) and NK (bottom) cells showing GFP marking. ( $n = 7, 11$ ). **d**, Expansion of GFP<sup>-</sup> and GFP<sup>+</sup> T and NK cells after tumour challenge. **e**, Tumour weight 3 weeks after challenge, in mice transplanted ( $n = 16$ ) or not ( $n = 3$ ). \*\*\*\* $P < 0.0001$  (unpaired  $t$ -test). **f**, NHEJ in the *IL2RG* gene on CD34<sup>+</sup> cells cultured *in vitro* and on their progeny sorted from the transplanted mice. **g**, TCR complexity score calculated on GFP<sup>+</sup> or GFP<sup>-</sup> T

cascade of  $\gamma$ -chain-coupled receptors (Fig. 4k and Extended Data Fig. 8). The targeted T cells displayed similar kinetics and extent of phosphorylation of STAT5 and AKT as their GFP<sup>-</sup> counterparts after stimulation with increasing IL-15 and IL-2 doses. Overall, these data prove functional reconstitution of the edited *IL2RG* gene, which supported lymphopoiesis and mature T-cell function indistinguishably from the wild-type allele.

### Specificity of *IL2RG* ZFNs on the HSC genome

We previously performed a genome-wide screening in K562 cells to identify potential off-target sites of the *IL2RG* ZFNs used in this study<sup>16</sup> and found a low rate of indels in up to 12 genomic loci bearing homology to the intended target site. We then determined whether these sites were also affected in the HSPCs treated here with ZFNs containing the same *IL2RG* DNA-binding domains but coupled to improved obligate heterodimeric FokI domains. We deep-sequenced the genomic regions encompassing the identified potential target sites on treated CD34<sup>+</sup> cells cultured *in vitro* and on human cells from the bone marrow of long-term engrafted NSG mice (Table 1 and Extended Data Fig. 9). The intended *IL2RG* target site showed 45–61% indel rate in the *in vitro* cultured cells and 20–43% in the *in vivo* engrafted cells. However, we detected indels only in two *in vitro* samples (at 0.17–0.7%) for the top two previously identified off-target sites, whereas from the *in vivo* samples

cells from transplanted mice. Human peripheral blood mononuclear cells (PBMCs) were used as positive control. \*\*\*\* $P < 0.0001$  (one-way ANOVA). **h**, *Ex vivo* growth of GFP<sup>+</sup> and GFP<sup>-</sup> T cells from the spleen of transplanted mice on stimulation ( $n = 4$ ). **i**, Division index of GFP<sup>+</sup> or GFP<sup>-</sup> T cells 7 days after PHA stimulation or co-culture with tumour cells at the indicated effector-to-target (E/T) ratios.  $P = NS$  (not significant) (unpaired  $t$ -test). T cells from healthy donor (HD) were used as controls. **j**, Southern blot (top), PCR (middle) and GFP qPCR (bottom) analyses showing targeted integration of the corrective *IL2RG* cDNA in sorted GFP<sup>+</sup> T cells from **h**. UT, untreated cells. **k**, Heat map showing changes in phosphorylation levels of STAT5 after the indicated time of exposure to decreasing amounts of IL-2 or IL-15, on T cells from **h**.  $P = NS$  (two-way ANOVA).

we found just one site with evidence of NHEJ (at 0.02%). Deep sequencing of all other putative off-target sites gave results not statistically different from the background error rate, which limits detection at 0.01% in our analysis (see Supplementary Information). The elimination of detectable off-target activity at some previously identified sites is consistent with the adoption of obligate heterodimeric FokI variants in this study, which would detarget activity from sites requiring ZFN homodimers. This analysis demonstrates a high specificity for the ZFNs used, evidenced by the 100-fold ratio between activities at the intended target site versus the top identified off-target site. It is possible, however, that additional off-target sites exist, which have not been identified by our previous screening.

### *IL2RG* gene correction in SCID-X1 HSPCs

We applied the optimized protocols developed for umbilical cord blood to CD34<sup>+</sup> cells from adult bone marrow and obtained an overall gene-targeting efficiency of  $6 \pm 0.5\%$  ( $n = 4$  donors) and a high rate of indels induced at either the AAVS1 or *IL2RG* ZFN target sites (Fig. 5a). Targeting was less efficient in the more primitive populations, although reaching similar values as those observed for umbilical cord blood cells. Xenotransplantation proved the long-term multilineage repopulation capacity of the targeted cells, with all transplanted NSG mice bearing GFP<sup>+</sup> cells at frequencies comparable to those observed with umbilical cord

**Table 1 | Target specificity of *IL2RG* ZFNs in HSPCs**

Nearest RefSeq gene	Intron/Exon	Homology (%)	ZFN dimer	NHEJ (%)					
				<i>In vitro</i>			Engrafted in mouse		
				A	B	G	B2	C0	E2
<i>IL2RG</i>	Exon	100	L_5_R	54.60	61.18	45.60	26.08	43.51	20.07
<i>SCARB1</i>	Outside	70.8	L_5_R	0.17	0.70	NS	NS	NS	NS
<i>SLC31A1</i>	Intron	75	R_5_L	0.61	NS	NS	NS	0.02	NS
<i>FAM133B</i>	Intron	66.7	R_6_R	NS	NS	NS	NS	NS	NS
<i>KIAA0528</i>	Intron	87.5	L_5_L	NS	NS	NS	NS	NS	NS
<i>SF3B1</i>	Outside	66.7	L_5_L	NS	NS	NS	NS	NS	NS
<i>A2BP1</i>	Outside	75	L_5_R	NS	NS	NS	NS	NS	NS
<i>ANKFY1</i>	Exon	87.5	L_3_R	NS	NS	NS	NS	NS	NS
<i>TRIM43</i>	Outside	91.7	L_4_L	NS	NS	NS	NS	NS	NS
<i>SEC16A</i>	Exon	70.8	R_6_L	NS	NS	NS	NS	NS	NS

The on-target and candidate off-target sites of *IL2RG* ZFNs were deep sequenced in the indicated progeny of treated CD34<sup>+</sup> cells. Intron/exon indicates whether the ZFN target site is within an exon, intron or intergenic (outside) near to the RefSeq gene indicated in the first column. Homology indicates the percentage of sequence identity to the *IL2RG* ZFN binding sites. ZFN dimer indicates the site can be bound by a homodimeric (LL/RR) or heterodimeric (LR/RL) ZFN pair; the number indicates the spacer length in base pairs between the ZFN-binding sites. NHEJ (%) indicates the percentage of indels on treated CD34<sup>+</sup> cells cultured *in vitro* (samples A, B and G) and on human cells harvested from the bone marrow of long-term engrafted NSG mice from Fig. 3 (mice B2, C0 and E2). For quantification see Supplementary Information and Extended Data Fig. 9. NS, not significant (Fisher's exact test for contingency data).

blood cells (Fig. 5b). On the basis of these results, we then tested our gene correction strategy on bone marrow CD34<sup>+</sup> cells from a symptomatic four-month-old SCID-X1 patient bearing a missense mutation in *IL2RG* exon 7 (c.865C>T; R289X). As expected for this mutation,

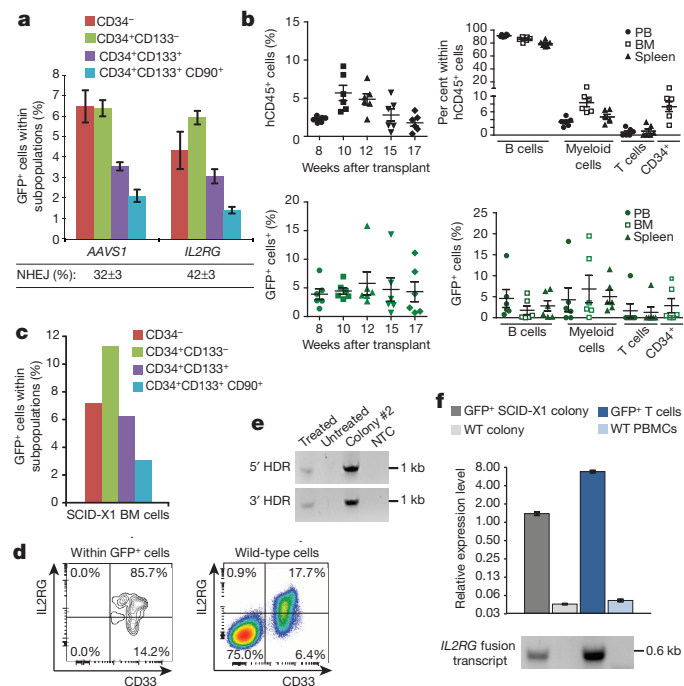
blood sampling or bone marrow harvest from the patient did not show any T or NK cells (Extended Data Fig. 10a, b). From 3% to 11% of the treated cell progeny became GFP<sup>+</sup>, depending on primitive versus committed progenitor status (Fig. 5c). CFC assays yielded three GFP<sup>+</sup> colonies out of ~100 scored (Extended Data Fig. 10c). Flow cytometry showed normal expression of the  $\gamma$ -chain protein in the myeloid progeny of the GFP<sup>+</sup> CFCs (Fig. 5d). Polymerase chain reaction (PCR) analyses of these colonies proved targeted integration into *IL2RG* leading to expression of the expected fusion transcript between the corrective cDNA and the upstream endogenous exons (Fig. 5e, f and Supplementary Information). Overall, these data show reconstitution of a functional *IL2RG* gene on targeted editing of a SCID-X1 allele in HSPCs.

## Discussion

Here we developed a strategy for targeted genome editing in human long-term repopulating HSCs and exploited it to insert transgenes into a genomic safe harbour or downstream of the promoter of an endogenous gene to reconstitute its functional expression. Using the latter approach, we demonstrate correction of the defective *IL2RG* gene in HSPCs from a subject with SCID-X1. As we obtained consistent results when targeting two different loci, we expect that our genome editing strategy can be used to target a variety of genomic sites. As the procedure uses a combination of IDLV infection and mRNA electroporation to deliver donor template and ZFNs, it has the potential to be applicable to a range of other genome editing tools.

We found that primitive cells are more sensitive than committed progenitors to the cytotoxicity of the gene targeting procedure and less proficient at performing HDR, probably because of their quiescence or slow cycling. These findings are consistent with reports showing delayed DNA repair and enhanced apoptosis after  $\gamma$ -radiation in human HSCs as compared to progenitors<sup>30</sup> and preferential repair of DNA DSBs by NHEJ in quiescent murine HSCs<sup>31</sup>. By delaying the time of treatment and exploiting compounds reported to support *ex vivo* maintenance and expansion of HSCs<sup>26–28</sup>, we were able to partially relieve the block to HDR. This effect is probably due to an increased transit through the S/G2 phases of the cell cycle, when HDR can occur, and, possibly, upregulation of its endogenous machinery. Other beneficial effects might be increased permissiveness to gene delivery, more efficient mRNA translation and reduced growth arrest and apoptosis in response to the gene targeting procedure. As improved procedures for *ex vivo* HSC expansion become available, they might increase the yield of gene targeted cells and allow their selection before *in vivo* administration. This would enable wide application of safe harbour sites, such as *AAVS1*, for robust expression of therapeutic transgenes<sup>22,23</sup>.

When applied to the functional correction of *IL2RG* in HSCs, we show that our strategy is compatible with normal development of the lymphoid lineages. *IL2RG*-edited lymphoid cells repopulated the mice and responded to  $\gamma$ -chain-dependent cytokines indistinguishably from



**Figure 5 | Targeted integration and *IL2RG* gene correction in bone-marrow-derived CD34<sup>+</sup> cells from healthy donors and a subject with SCID-X1.** **a**, Top: GFP<sup>+</sup> cells within the indicated subpopulations derived from bone marrow CD34<sup>+</sup> cells of adult healthy donors, treated for targeted integration according to the best performing protocol from Fig. 3. Bottom: NHEJ at the ZFN target site on total cells. Means  $\pm$  s.e.m., ( $n = 10$ , 3 from 4, 3 donors for *AAVS1* or *IL2RG*, respectively). **b**, Top left: human cells in peripheral blood of NSG mice transplanted with cells from **a**. Top right: percentages of the indicated lineages within human cells 15 weeks after transplant. Bottom: GFP<sup>+</sup> cells within the indicated populations. **c**, GFP<sup>+</sup> cells measured as in **a** in bone marrow CD34<sup>+</sup> cells from a subject with SCID-X1 treated for *IL2RG* gene correction. **d**, *IL2RG* expression in myeloid (CD33<sup>+</sup>) cells from a GFP<sup>+</sup> colony from the cells treated in **c** or from pooled wild-type colonies. **e**, PCR analysis for targeted integration into *IL2RG* of the corrective cDNA on cells from **c** and **d**. **f**, Expression of the fusion transcript bearing the corrective *IL2RG* cDNA measured by qPCR (top) or RT-PCR (bottom) on cDNA from a GFP<sup>+</sup> SCID-X1 myeloid colony. *IL2RG* targeted T cells from engrafted mice analysed in Fig. 4j were used as positive control, whereas a myeloid colony from wild-type bone marrow cells (WT) and PBMCs were used as negative controls.



their unedited counterparts. On the contrary, lymphoid cells carrying NHEJ-mediated *IL2RG* inactivation were counter-selected in the mice, phenocopying the SCID-X1 disease<sup>1</sup>.

This disease may be suitable for the clinical translation of targeted gene correction, as previous clinical studies have demonstrated the potential efficacy but also the risks of HSC gene therapy using early generation vectors<sup>7,8</sup>. Whereas our strategy should abrogate the risk of insertional mutagenesis and the concern for unregulated transgene expression, it remains to be established whether a potentially limiting amount of targeted HSPCs would enable effective and safe correction of SCID-X1 in humans (even though both HSCs and progenitors, which are more efficiently targeted, would help repopulating the absent lymphoid lineages). Another concern is with the potential off-target activity of ZFNs, although our analysis showed a high specificity for the *IL2RG* ZFNs used in this study. Overall, we envisage a fairly straightforward path to clinical translation of our strategy, considering that good manufacturing practice of lentiviral vector and mRNA electroporation are already established. Moreover, infusion of autologous gene-targeted cells could be performed without pharmacological conditioning, taking advantage of the selective advantage of the corrected cells.

## METHODS SUMMARY

Donor IDLV for HDR were generated as described. ZFNs targeting intron 1 of *PPP1R12C* or exon 5 of *IL2RG* (refs 16, 22) were expressed by mRNA electroporation. CD34<sup>+</sup> cells from human umbilical cord blood or bone marrow were used on approval by the San Raffaele Hospital Bioethical Committee, stimulated in serum-free medium with early acting cytokines, infected with IDLVs at a multiplicity of infection (MOI) 100–500, and then electroporated with 175 µg ml<sup>-1</sup> ZFNs encoding mRNAs. Targeted integration was assessed by PCR and Southern blot while ZFN activity was determined by Cell1 assay and deep sequencing of genomic target sites. The treated CD34<sup>+</sup> cells at day 4 of culture were infused intravenously into sublethally irradiated 8–11-week-old NOD-SCID-*Il2rg*<sup>-/-</sup> (NSG) mice. To expand human T and NK cells, 4 × 10<sup>6</sup> MDA3-MB231 tumour cells expressing human IL-7, IL-15 and GM-CSF were implanted orthotopically in the mammary fat pad of NSG mice 14 weeks after CD34<sup>+</sup> cells transplantation<sup>29</sup>. Functional assays on *IL2RG*-edited T cells collected from transplanted mice were carried out after stimulation with beads conjugated to anti-human CD3 and CD28 antibodies and sorting for GFP expression.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 August 2013; accepted 29 April 2014.

Published online 28 May 2014.

- Mukherjee, S. & Thrasher, A. J. Gene therapy for PIDs: progress, pitfalls and prospects. *Gene* **525**, 174–181 (2013).
- Aiuti, A. *et al.* Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341**, 1233151 (2013).
- Cavazzana-Calvo, M. *et al.* Transfusion independence and *HMG2* activation after gene therapy of human  $\beta$ -thalassaemia. *Nature* **467**, 318–322 (2010).
- Cartier, N. *et al.* Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* **326**, 818–823 (2009).
- Biffi, A. *et al.* Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* **341**, 1233158 (2013).
- Naldini, L. *Ex vivo* gene transfer and correction for cell-based therapies. *Nature Rev. Genet.* **12**, 301–315 (2011).
- Braun, C. J. *et al.* Gene therapy for Wiskott-Aldrich syndrome—long-term efficacy and genotoxicity. *Science Transl. Med.* **6**, 227ra33 (2014).
- Cavazza, A., Moiani, A. & Mavilio, F. Mechanisms of retroviral integration and mutagenesis. *Hum. Gene Ther.* **24**, 119–131 (2013).
- Woods, N. B., Bottero, V., Schmidt, M., von Kalle, C. & Verma, I. M. Gene therapy: therapeutic gene causing lymphoma. *Nature* **440**, 1123 (2006).
- Gentner, B. *et al.* Identification of hematopoietic stem cell-specific miRNAs enables gene therapy of globoid cell leukodystrophy. *Sci. Transl. Med.* **2**, 58ra84 (2010).

- Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Natl. Rev.* **11**, 636–646 (2010).
- Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nature Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
- Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnol.* **32**, 347–355 (2014).
- Lombardo, A. *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotechnol.* **25**, 1298–1306 (2007).
- Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
- Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nature Biotechnol.* **29**, 816–823 (2011).
- Mussolino, C. *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39**, 9283–9293 (2011).
- Ciccica, A. & Elledge, S. J. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40**, 179–204 (2010).
- Tebas, P. *et al.* Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **370**, 901–910 (2014).
- Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 *in vivo*. *Nature Biotechnol.* **28**, 839–847 (2010).
- Provati, E. *et al.* Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nature Med.* **18**, 807–815 (2012).
- Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nature Methods* **8**, 861–869 (2011).
- Zou, J. *et al.* Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. *Blood* **117**, 5561–5572 (2011).
- Li, H. *et al.* *In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221 (2011).
- Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: a human perspective. *Cell Stem Cell* **10**, 120–136 (2012).
- Boitano, A. E. *et al.* Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science* **329**, 1345–1348 (2010).
- North, T. E. *et al.* Prostaglandin E2 regulates vertebrate hematopoietic stem cell homeostasis. *Nature* **447**, 1007–1011 (2007).
- Goessling, W. *et al.* Prostaglandin E2 enhances human cord blood stem cell xenotransplants and shows long-term safety in preclinical nonhuman primate transplant models. *Cell Stem Cell* **8**, 445–458 (2011).
- Escobar, G. *et al.* Genetic engineering of hematopoiesis for targeted IFN- $\alpha$  delivery inhibits breast cancer progression. *Sci. Transl. Med.* **6**, 217ra213 (2014).
- Milyavsky, M. *et al.* A distinctive DNA damage response in human hematopoietic stem cells reveals an apoptosis-independent role for p53 in self-renewal. *Cell Stem Cell* **7**, 186–197 (2010).
- Mohr, M. *et al.* Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell* **7**, 174–185 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank D. Weissman for advice on mRNA production and the whole Naldini laboratory for discussion, F. Benedicenti for help with MiSeq sequencing, L. Sergi, T. Plati, V. Valtolina, B. Camisa and A. Ranghetti for technical help. SR1 was provided by T. Boitano and M. Cooke under an MTA with the Genomics Institute of the Novartis Research Foundation. This work was supported by grants to L.N. from Telethon (TIGET grant D2) EU (FP7 222878 PERSIST, FP7 601958 SUPERSIST, ERC Advanced Grant 249845 TARGETINGTHERAPY) and the Italian Ministry of Health.

**Author Contributions** P.G. designed experiments, performed research, interpreted data and wrote the manuscript. G.S. and G.E. performed research and interpreted data. T.D.T. performed mRNA production. C.F. characterized the corrective cDNA. A.C. and E.M. performed bioinformatics analysis of ZFN specificity. R.M. and D.M. developed the NSG human tumour rejection model. C.B. contributed to the T-cell studies. M.v.d.B. provided SCID-X1 patient cells. M.C.H. and P.D.G. provided ZFNs, interpreted data and edited the manuscript. B.G. set up culture conditions for HSC maintenance. A.L. and L.N. designed and supervised research, interpreted data and wrote the manuscript. L.N. coordinated the study. G.S. and G.E. contributed equally to this work. A.L. and L.N. share senior authorship.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.N. ([naldini.luigi@hsr.it](mailto:naldini.luigi@hsr.it)).

## METHODS

**Vectors and zinc finger nucleases.** Homology-directed repair donor templates were generated from HIV-derived, third-generation self-inactivating transfer constructs. IDLV stocks were prepared as previously described<sup>14</sup> and titred by a qPCR designed to discriminate the reverse-transcribed vector genome from plasmid carried over from transient transfection<sup>32</sup>. Sequence and maps of *AAVS1-PGK.GFP* were previously reported<sup>22</sup>, whereas *IL2RG-cDNA.PGK.GFP* is described in detail elsewhere (Firrito *et al.*, manuscript in preparation). ZFNs that target intron 1 of *PPP1R12C* or exon 5 of *IL2RG* were previously described<sup>14,16,22,33</sup>. The latter pair was modified to contain high-fidelity obligate heterodimeric FokI variants<sup>34</sup>. Both pairs of ZFNs were transiently expressed as mRNAs. Plasmid templates for ZFNs mRNA production (described in Extended Data Fig. 1b) were linearized and purified by phenol/chloroform extraction followed by ethanol DNA precipitation. 2 µg per reaction of linearized plasmid template was *in vitro* transcribed at 37 °C for 2 h using T7 RNA polymerase and 7.5 mM nucleotide triphosphates (MEGAscript kit; Ambion). Cap0 mRNAs were generated by supplementing the reactions with 6 mM m<sup>7</sup>(3'-O-methyl)-G(5')ppp(5')G, a non-reversible cap analogue (ARCA, New England Biolabs) and lowering the concentration of GTP to 1.5 mM. After TURBO DNase treatment (4 U per reaction, 1 h at 37 °C), mRNAs were poly(A) tailed with *E. coli* Poly(A) Polymerase (8 U per reaction) for 1 h at 37 °C (PolyA tailing kit; Ambion), yielding ≥150 nt polyA. Transcripts were purified by the RNeasy Plus Mini Kit (Qiagen). All RNA samples were analysed by denaturing agarose gel electrophoresis for quality assurance.

**In vitro culture and assays on human cord blood or bone-marrow-derived CD34<sup>+</sup> cells.** CD34<sup>+</sup> cells were either freshly purified from human cord blood after obtaining informed consent and on approval by the San Raffaele Hospital Bioethical Committee, or purchased frozen from Lonza. 10<sup>6</sup> CD34<sup>+</sup> cells per ml were stimulated in serum-free StemSpan medium (StemCell Technologies) supplemented with penicillin, streptomycin and human early-acting cytokines (for cord-blood-derived cells: stem cell factor (SCF) 100 ng ml<sup>-1</sup>, Flt3 ligand (Flt3-L) 100 ng ml<sup>-1</sup>, thrombopoietin (TPO) 20 ng ml<sup>-1</sup>, and interleukin 6 (IL-6) 20 ng ml<sup>-1</sup>; for bone-marrow-derived cells: SCF 300 ng ml<sup>-1</sup>, Flt3-L 300 ng ml<sup>-1</sup>, TPO 100 ng ml<sup>-1</sup>, and IL-6 60 ng ml<sup>-1</sup>; all purchased from Peprotech) for 24 or 48 h and then infected with IDLVs at a multiplicity of infection (MOI) 100–500. The following day the cells were electroporated with 175 µg ml<sup>-1</sup> ZFNs encoding mRNAs (P3 Primary Cell 4D-Nucleofector X Kit, program EO-100; Lonza). For some experiments, the following drugs were supplemented to the culture media: 1 µM SR1 (provided by T. Boitano and M. Cooke, GNF) added at every medium change, and 10 µM dmPGE2 (Cayman) added at the beginning of the culture, 1 h before and just after electroporation. For CFC assays, 800 cells per plate were seeded one day after electroporation in methylcellulose-based medium (MethoCult H4434, StemCell Technologies). Two weeks after plating, colonies were counted and identified according to morphological criteria.

**Flow cytometry.** For immunophenotypic analysis of CD34<sup>+</sup> cells and their progeny (performed on FACSCanto II; BD Pharmingen), we used the antibodies reported in Supplementary Information. Single-stained and FMO-stained cells were used as controls. For quantitative flow cytometry we used Flow-count Fluorospheres (Beckman Coulter) according to the manufacturers' instructions. Apoptosis analysis was performed on CD34<sup>+</sup> cells one day after electroporation using peripheral-blood-conjugated Annexin V (Biolegend) and Apoptosis Detection kit with 7-Aminoactinomycin D (7AAD, BD Pharmingen) according to the manufacturers' instructions. Percentages of live (7AAD<sup>-</sup>, AnnexinV<sup>-</sup>), early apoptotic (7AAD<sup>-</sup>, AnnexinV<sup>+</sup>), late apoptotic (7AAD<sup>+</sup>, AnnexinV<sup>+</sup>) and necrotic (7AAD<sup>+</sup>, AnnexinV<sup>-</sup>) cells are reported. Cell sorting was performed using MoFlo XDP Cell Sorter (Beckman Coulter).

**Molecular analyses.** For molecular analyses, genomic DNA was isolated with DNeasy Blood & Tissue Kit or QIAamp DNA Micro Kit (QIAGEN). Extraction of genomic DNA from colonies in CFC assays was performed with Lysis Buffer as previously described<sup>2</sup>. NHEJ in *AAVS1* locus or *IL2RG* gene was detected by the mismatch selective Cel1 assay as previously described<sup>14</sup>. Primers for PCR amplifications to detect targeted integration or for the Cel1 assay are indicated in Supplementary Information. PCR amplicons were resolved on agarose gel and visualized by ethidium bromide staining. For Southern blot analyses, genomic DNA was extracted with Blood & Cell Culture DNA Midi Kit (QIAGEN) and digested using restriction enzymes (BglI for *AAVS1* locus and BspHI for *IL2RG*) and probes (see Supplementary Information) both located outside of the homology regions included in the vectors. Matched DNA amounts were separated on 1% agarose, transferred to a nylon membrane and probed with <sup>32</sup>P-radiolabelled sequences indicated in Supplementary Information. Membranes were exposed in a Storage Phosphor Screen. For qPCR analysis, 200 ng of genomic DNA were analysed using primers and probes complementary to a vector backbone sequence (Primer Binding Site), the GFP sequence and human *TERT*, the latter amplification used as normalizer, as previously described<sup>14</sup>. For gene expression

analysis on the *SCID-X1* gene corrected colony, mRNA was extracted using the RNeasy Micro Kit (QIAGEN) and cDNA was synthesized using the SuperScript VILO cDNA Synthesis Kit (Invitrogen). The resulting cDNA was amplified before qPCR by Taqman PreAmp Master Mix Kit (Applied Biosystems) according to the manufacturers' instructions. Gene expression was performed in triplicate with a TaqMan Gene Expression assay specific for the recoded exon 7 of the *IL2RG* gene (Applied Biosystems; see Supplementary Information) in a 7900HT real-time PCR thermal cycler. The relative expression level of the recoded *IL2RG* gene was calculated by the  $\Delta\Delta C_t$  method and represented as fold change relative to the housekeeping gene control (*HPRT*), as previously described<sup>22</sup>.

**Mice transplantation and analysis.** For the *in vivo* studies, 8- to 11-week-old NOD-SCID-*Il2rg*<sup>-/-</sup> (NSG) mice were purchased by Jackson laboratory. The experimental protocol was approved by the Institutional Animal Care and Use Committee of the San Raffaele Scientific Institute. At day 4 of culture, 3 × 10<sup>5</sup> gene targeted cord-blood-derived CD34<sup>+</sup> cells (or 7.5 × 10<sup>5</sup> bone-marrow-derived cells) were infused intravenously into the mice after sub-lethal irradiation (200 cGy). Sample size was determined by the total number of available treated cells. Mice were attributed to each experimental group randomly. MDA3 human mammary carcinoma cell line was obtained by stable transduction of MDA-MB 231 cells with lentiviral vectors expressing the human cytokines GM-CSF, IL-7 and IL-15 from the PGK promoter, as previously described<sup>29</sup>. 4 × 10<sup>6</sup> MDA3 cells were implanted orthotopically in the mammary fat pad of NSG mice 14 weeks after CD34<sup>+</sup> cell transplantation or in age-matched untransplanted NSG mice. Human CD45<sup>+</sup> engraftment was followed by serial collections of blood from the mouse tail and, at the end of the experiment (12–23 weeks after transplantation), bone marrow and spleen were harvested and analysed for lineage composition and GFP content (see Supplementary Information for gating strategy).

**T lymphocyte analyses.** Human T cells were enriched from splenocytes harvested from the transplanted mice using magnetic beads conjugated to anti-human CD3 and CD28 antibodies (Dynabeads human T-activator CD3/CD28; Invitrogen), following the manufacturers' instructions, and grown in Iscove's Modified Dulbecco's Media (IMDM) (GIBCO-BRL) supplemented with penicillin, streptomycin, 10% FBS and 5 ng ml<sup>-1</sup> each of IL-7 and IL-15 (PeproTech)<sup>21</sup>.

For TCR V-β repertoire analysis, mRNA was extracted from the expanded T cells using the RNeasy Mini Kit (QIAGEN) and cDNA was synthesized using the SuperScript VILO cDNA Synthesis Kit (Invitrogen). Multiplex PCRs optimized from a previous work<sup>35</sup> were carried out on cDNA using V-β primers specific for 4 or 5 different families and a single FAM-labelled C-β primer. PCR products were fractionated on 6% polyacrylamide gel, visualized on Molecular Dynamics Typhoon 9410 (Amersham Biosciences) and analysed using ImageQuant TL 7.0 (Amersham Biosciences). V-β complexity was determined by counting the number of distinct peaks and graded on a score of 0–8 (ref. 36). The overall TCR complexity score was determined by summing up all 23 individual TCR V-β family-specific scores.

To analyse phosphorylation of downstream effectors of the IL-2RG pathway, T cells were starved overnight at 37 °C in IMDM without cytokines and then stimulated with IL-2 (1,000 IU ml<sup>-1</sup>, 100 IU ml<sup>-1</sup>, 10 IU ml<sup>-1</sup>; purchased from Novartis) or IL-15 (10 ng ml<sup>-1</sup>, 5 ng ml<sup>-1</sup>, 1 ng ml<sup>-1</sup>) at 37 °C for increasing times. Cells were then fixed in PBS 2% paraformaldehyde (PFA) for 10 min at 37 °C, and after washing in PBS 0.1% BSA (3 times), they were permeabilized with ice-cold absolute methanol for 7 min on ice. After 60 min incubation of each time point of cytokine stimulation with different dilutions of Pacific blue succinimidyl ester (PBSE) (Life Technologies), cells were washed, pooled and stained for flow cytometry.

For proliferation assay, 10<sup>5</sup> T cells were labelled with Cell Proliferation Dye eFluor 670 (eBioscience) according to the manufacturer's instructions. Labelled T cells were co-cultured in IMDM supplemented with penicillin, streptomycin, 10% FBS and 5 ng ml<sup>-1</sup> each of IL-7 and IL-15, with different dilutions of MDA-MB 231 cells previously irradiated at 10,000 rad or stimulated for 3 days with PHA (2 µg ml<sup>-1</sup>). After 7 days of culture, cells were analysed by flow cytometry. Division index was calculated according to FlowJo software rules.

For IFN-γ release assay, T cells were stimulated at 37 °C for 6 h with PMA (50 ng ml<sup>-1</sup>) and ionomycin (1 µg ml<sup>-1</sup>) in presence of 2 µl per ml of culture of BD Golgi Plug (BD Pharmingen). Cells were then fixed and permeabilized using the BD Cytofix/Cytoperm Kit (BD Pharmingen) and stained for flow cytometry.

Elispot assay for IFN-γ release was performed as previously described<sup>21,29</sup>. When the number of measured spots was above the detection limit of the plate reader (EliExpert, A.E.L.VIS.), it was arbitrarily set to 500.

**Deep sequencing of potential *IL2RG* ZFN off-target loci.** Genomic DNA from ZFN-treated CD34<sup>+</sup> cells or their progeny harvested from transplanted mice was amplified using the REPLI-g Mini Kit (QIAGEN) and the top-ranking candidate off-target genomic loci from our previous study<sup>16</sup> amplified by PCR, generating amplicons of 389 ± 20 bp surrounding the potential ZFN binding site. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter) and adaptors were added by TruSeq DNA LT Sample Prep Kit (Illumina). To build an

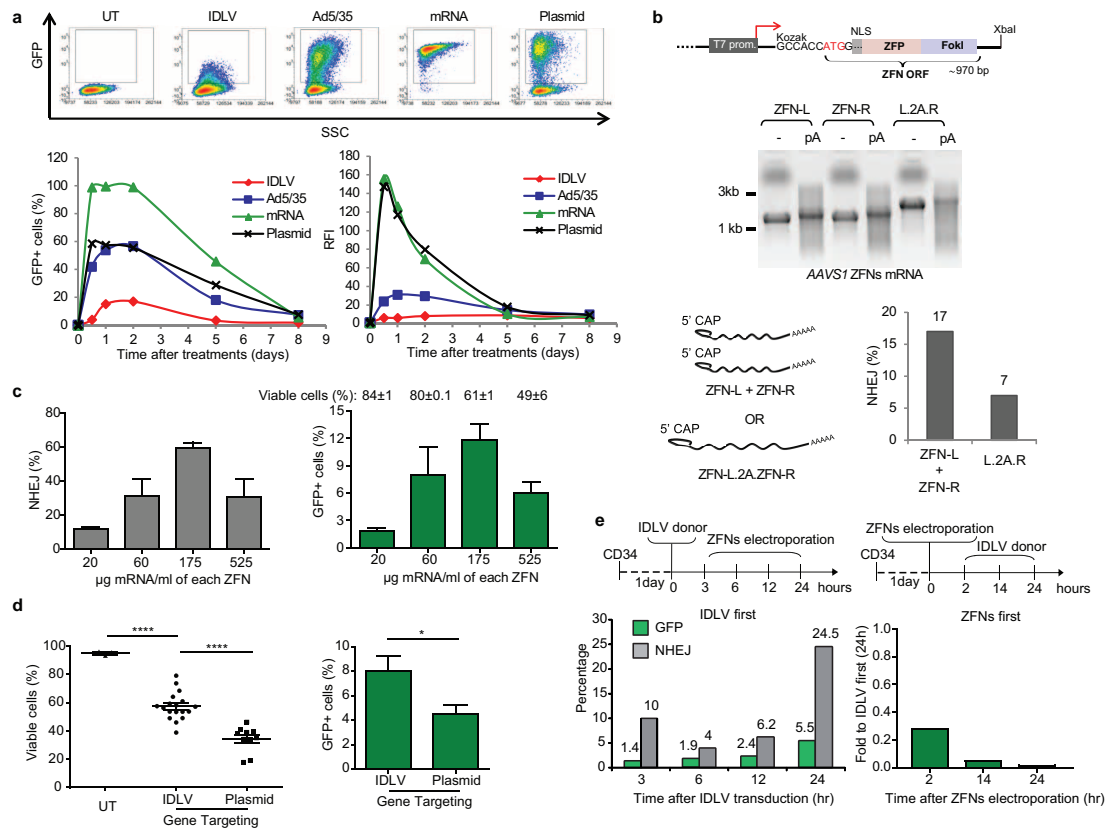


equimolar library, PCR products were quantified with KAPA Library Quantification Kit for Illumina sequencing platforms (KAPABIOSYSTEMS) on C1000 Thermal Cycler (BIO-RAD) and sequenced on MiSeq Illumina Platform using MiSeq Reagent v.3 (Illumina). Raw paired-end reads were joined with Fastq-Join program from the EA-Utils NGS suite (<http://code.google.com/p/ea-utils/>) and aligned to the specific genomic target sequences using Burrows–Wheeler Alignment Tool with maximal exact match version, BWA-MEM<sup>37</sup>. Alignments were evaluated and filtered using SAMtools<sup>38</sup>, Picard (<http://picard.sourceforge.net>) and BAMtools<sup>39</sup>. Sequences with only primary alignments with quality >15 were kept for further analysis. Deletions and insertions (indels) were quantified by a custom pipeline based on Python (<http://www.python.org>, version 2.7.6) and the PySAM library (<https://code.google.com/p/pysam>, version 0.7.5). Sequences with indels of  $\geq 1$  bp located within a region encompassing the spacer +5 bp on each side were considered as ZFN-induced genome modifications. Coverage statistics were computed by the binomial distribution online calculator (<http://www.vassarstats.net/binomialX.html>). Fisher exact test were computed with the SciPy Python package (<http://www.scipy.org>, version 0.9.0) within the ‘stats’ library. Multiple sequence alignment for indels visualization and plot was performed with ClustalW2<sup>40</sup> and MView (<http://bio-mview.sourceforge.net/>).

**SCID-X1 cells.** Peripheral blood and bone-marrow samples from a subject with SCID-X1 were obtained according to the guidelines of the Medical Ethics Committee of the Erasmus MC, University Medical Center Rotterdam, the Netherlands.

**Statistical analyses.** Statistical analyses were performed by unpaired Student’s *t*-test for pairwise comparison or one-way or two-way analysis of variance (ANOVA) with Bonferroni’s multiple comparison post-test for three or more groups, as indicated. Values are expressed as mean  $\pm$  standard error of the mean (s.e.m.). Per cent values were transformed into a log-odds scale ( $\log(\% \times (100 - \%))$ ) to perform statistical analyses.

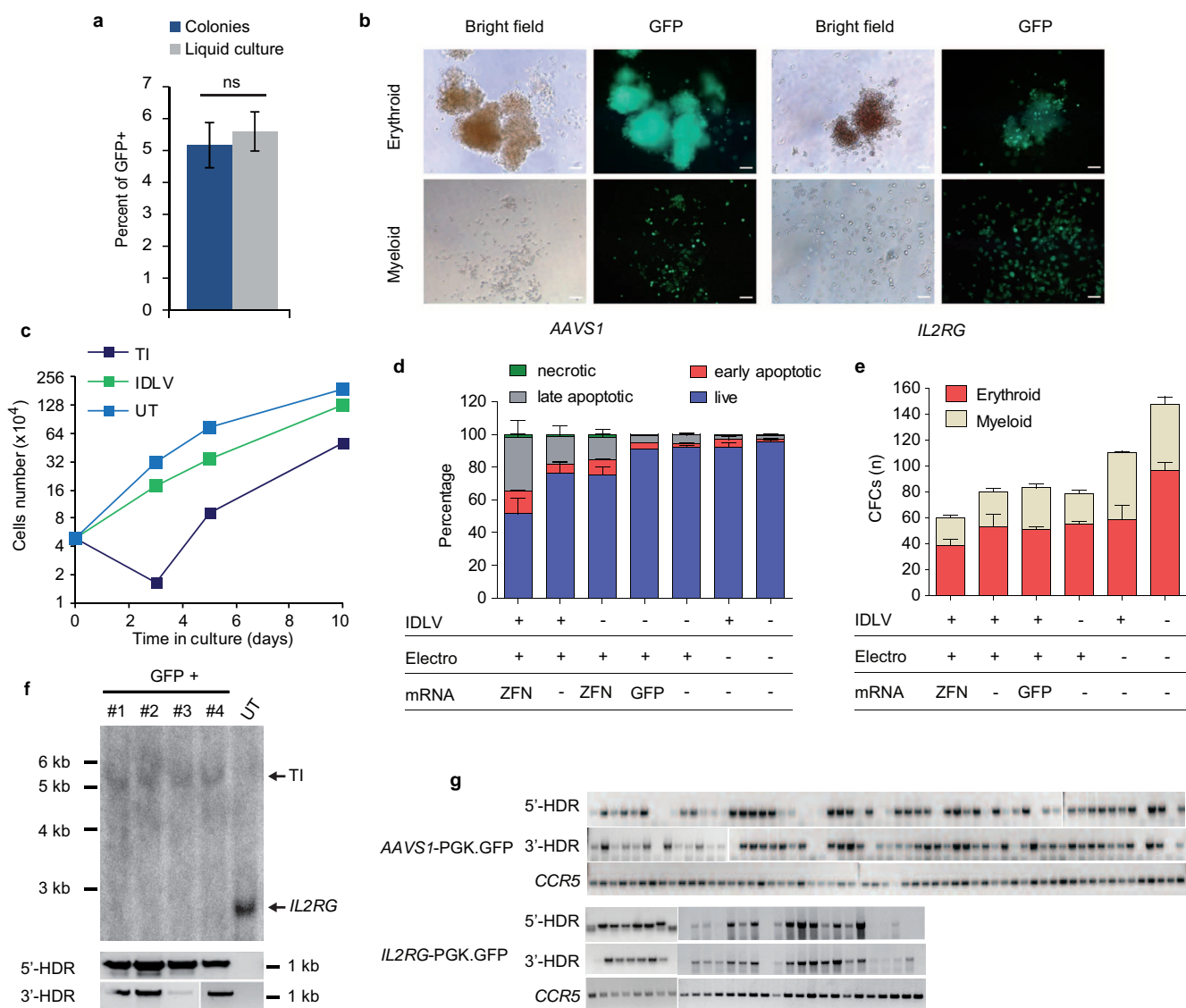
32. Mátrai, J. *et al.* Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatology* **53**, 1696–1707 (2011).
33. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature Biotechnol.* **27**, 851–857 (2009).
34. Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnol.* **25**, 778–785 (2007).
35. Akatsuka, Y., Martin, E. G., Madonik, A., Barsoukov, A. A. & Hansen, J. A. Rapid screening of T-cell receptor (TCR) variable gene usage by multiplex PCR: application for assessment of clonal composition. *Tissue Antigens* **53**, 122–134 (1999).
36. Wu, C. J. *et al.* Reconstitution of T-cell receptor repertoire diversity following T-cell depleted allogeneic bone marrow transplantation is related to hematopoietic chimerism. *Blood* **95**, 352–359 (2000).
37. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
38. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
39. Barnett, D. W., Garrison, E. K., Quinlan, A. R., Stromberg, M. P. & Marth, G. T. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691–1692 (2011).
40. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
41. Harris, D. T., Badowski, M., Balamurugan, A. & Yang, O. O. Long-term human immune system reconstitution in non-obese diabetic (NOD)-Rag (-)- $\gamma$  chain (-) (NRG) mice is similar but not identical to the original stem cell donor. *Clin. Exp. Immunol.* **174**, 402–413 (2013).
42. Gattinoni, L. *et al.* A human memory T cell subset with stem cell-like properties. *Nature Med.* **17**, 1290–1297 (2011).
43. Cieri, N. *et al.* IL-7 and IL-15 instruct the generation of human memory stem T cells from naive precursors. *Blood* **121**, 573–584 (2013).



**Extended Data Figure 1 | Optimization of gene targeting protocol in CD34<sup>+</sup> cells.** We optimized the delivery platform, dose and timing of ZFNs and HDR donor template administration. **a**, Performance of different gene delivery platforms. Cord blood (CB) CD34<sup>+</sup> cells were pre-stimulated with early acting cytokines for 24 h and transduced with GFP-encoding IDLV (MOI  $5 \times 10^2$ ) or adenoviral vector serotype 5/35 (MOI  $5 \times 10^3$ ), or electroporated with GFP-expressing mRNA (500  $\mu\text{g ml}^{-1}$ ) or plasmid DNA (25  $\mu\text{g ml}^{-1}$ ). The cells were analysed by flow cytometry at the indicated days after the procedure. Top: representative density plots of GFP expression 24 h post treatment. SSC, side scatter. Bottom: kinetics of transgene expression measured as a percentage of GFP<sup>+</sup> cells (left) and relative GFP fluorescence intensity (RFI, measured as the ratio between the mean fluorescence intensity of the treated cells at each time point to the untreated cells) in arbitrary units (right). UT, untreated cells. mRNA electroporation outperformed all approaches tested in terms of frequency of transfected cells and protein expression level. Note that although IDLV infects the majority of cells in these conditions its expression is constrained by the unintegrated nature<sup>32</sup>. Because mRNA transfection drives a robust but short-lived spike of expression, it appeared best suited for ZFN delivery, allowing proficient activity of the nucleases at the genomic target site while avoiding prolonged exposure. **b**, Top: schematic representation (not in scale) of a plasmid DNA template used for *in vitro* mRNA transcription with the T7 promoter, the Kozak sequence and the XbaI restriction enzyme used for the plasmid linearization depicted. The protein domains of a ZFN are shown within the open reading frame (ORF). NLS, nuclear localization signal; ZFP, zinc finger protein; FokI, FokI nuclease domain. Middle: representative denaturing gel electrophoresis of *in vitro* transcribed mRNAs encoding for the pair of ZFNs specific for AAVS1, before (–) and after enzymatic polyadenylation (pA). The ZFN mRNAs were produced either as two separated transcripts (ZFN-L and ZFN-R) or as a single construct coding for both ZFNs linked by a Tav.2A self-cleavage peptide sequence (ZFN-L.2A.ZFN-R; bottom left). Bottom right: cord blood CD34<sup>+</sup> cells were electroporated either with the two separate transcripts or with the single mRNA co-expressing both ZFNs. ZFN activity was measured on treated cells as percentage of NHEJ detected at the ZFN target site by Cell assay 10 days after electroporation. **c**, Dose-response for AAVS1 targeting ZFN mRNAs in CD34<sup>+</sup> cells. Cord blood CD34<sup>+</sup> cells were transduced with integrase-defective lentiviral vector (IDLV)<sup>14</sup> bearing homology to the AAVS1 locus and expressing GFP, and then electroporated with the indicated escalating doses of AAVS1 ZFN mRNAs. ZFN activity was scored by measuring the extent of NHEJ-mediated repair at

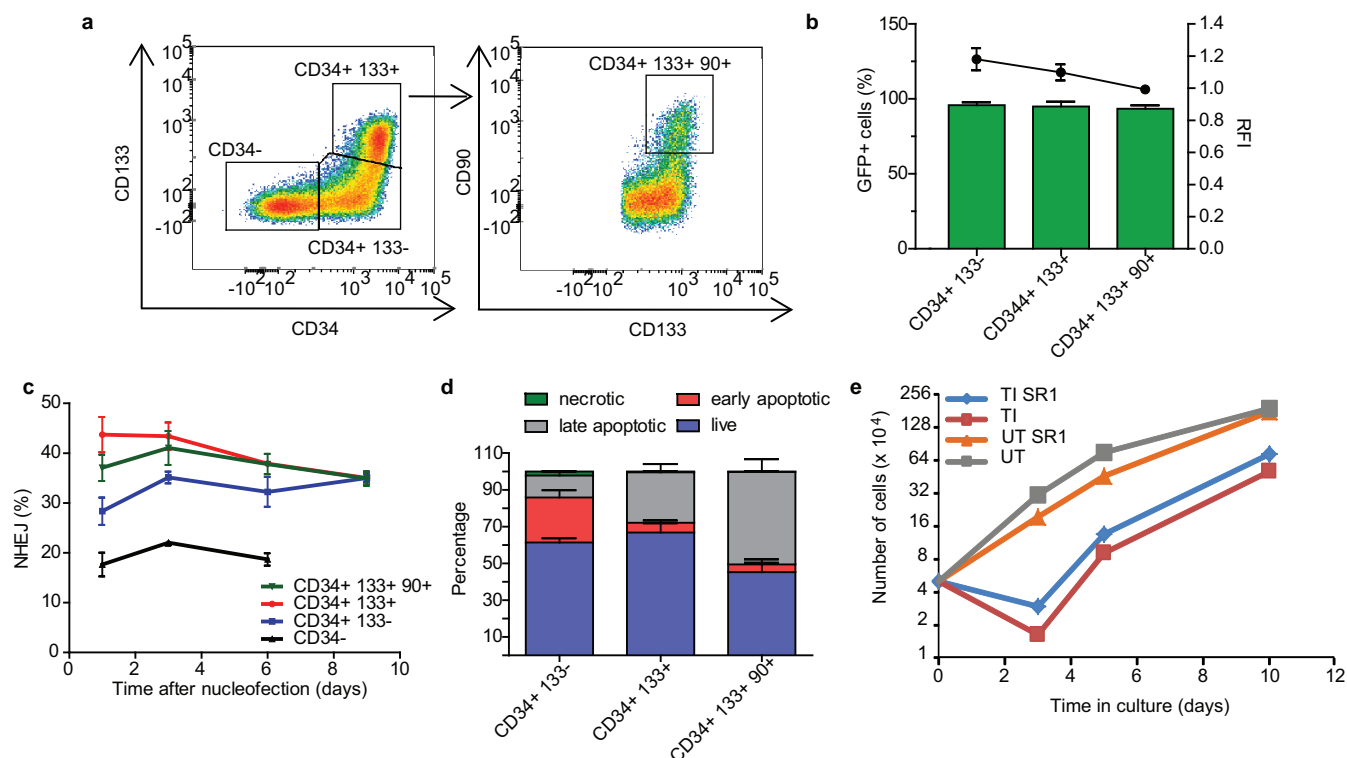
their genomic target site, and HDR was scored by the frequency of GFP<sup>+</sup> cells obtained in liquid culture. Left: NHEJ measured by Cell assay at day 10 after electroporation for the indicated dose of mRNA. Means  $\pm$  s.e.m. ( $n = 3$ ). Right: percentages of GFP<sup>+</sup> cells by flow cytometry 3 days after treatment. The percentages of viable cells (indicated on top of the histogram) were calculated as percentages of 7AAD negative cells gated on singlets. A dose-dependent increase in the percentage of NHEJ and GFP<sup>+</sup> cells was observed for the first three mRNA doses, whereas the highest dose caused a significant reduction in the number of viable cells, and a reduction in the efficiency of gene targeting. Based on these data, we selected the dose of 175  $\mu\text{g ml}^{-1}$  RNA to perform all further experiments. **d**, Choice of delivery platform for the HDR donor template. Cord blood CD34<sup>+</sup> cells were either transduced with the AAVS1 donor IDLV and electroporated with the cognate ZFNs mRNAs, or co-electroporated with AAVS1 donor plasmid DNA and ZFN mRNAs. Left: cell viability measured by flow cytometry 3 days after electroporation, comparing untreated cells (UT) and gene-targeted cells using IDLV or plasmid as donor templates. \*\*\*\* $P < 0.0001$  (one-way ANOVA with Bonferroni's multiple comparison post-test). Right: percentage of GFP<sup>+</sup> cells using either donor templates. Means  $\pm$  s.e.m. (UT,  $n = 3$ ; IDLV,  $n = 18$ ; plasmid,  $n = 10$ ). \* $P < 0.05$  (unpaired *t*-test). IDLV infection outperformed plasmid DNA electroporation in terms of the frequency of GFP<sup>+</sup> cells and cell viability, consistent with our previous findings in other primary cell types<sup>14,22</sup>. **e**, Schedule optimization for ZFNs and donor template delivery. After one day of pre-stimulation, cord blood CD34<sup>+</sup> cells were first transduced with the AAVS1 donor IDLV and then electroporated at the indicated hours post-infection with ZFN mRNAs (left) or, on the contrary, first electroporated with ZFN mRNAs and then transduced with IDLV (right). The time lines of the experiments are shown on top of the histograms. The percentages of GFP<sup>+</sup> cells measured by flow cytometry 3 days after treatment and NHEJ measured by Cell assay 10 days after treatment are shown on bottom left. Bottom right: the percentage of GFP<sup>+</sup> cells is expressed as fold to the percentage achieved in the same experiment with the best strategy on the left. The highest frequency of GFP<sup>+</sup> cells was obtained by combining IDLV-based donor template delivery 24 h before ZFN mRNA electroporation. Sequential exposure to the two delivery platforms avoids competition for cell entry and minimizes mutual interference, probably due to activation of innate responses to exogenous nucleic acids or the timing of peak ZFN expression relative to IDLV reverse transcription and nuclear import.





**Extended Data Figure 2 | Impact on cell viability and specificity of integration in CD34<sup>+</sup> cells treated for targeted integration.** **a**, Percentage of GFP<sup>+</sup> cells measured in liquid culture 3 days after treatment as in Fig. 1b for targeted integration into *AAVS1* or *IL2RG* and for the corresponding GFP<sup>+</sup> colonies counted in CFC assays 2 weeks after plating. Means  $\pm$  s.e.m. ( $n = 7$  cord blood donors). ns, not significant (unpaired *t*-test). **b**, Representative bright-field and fluorescence microscopy images of GFP<sup>+</sup> erythroid and myeloid colonies. Scale bar, 0.5 mm. **c–e**, The impact of the gene-targeting procedure on the viability, proliferation and clonogenic output of the CD34<sup>+</sup> cells was analysed. **c**, Representative growth curves of CD34<sup>+</sup> cells treated for targeted integration or transduced with IDLV only or untreated (UT). **d**, Apoptosis analysis performed 24 h after treatment on cells in liquid culture and **e**, number of CFCs plated one day after the indicated treatments.

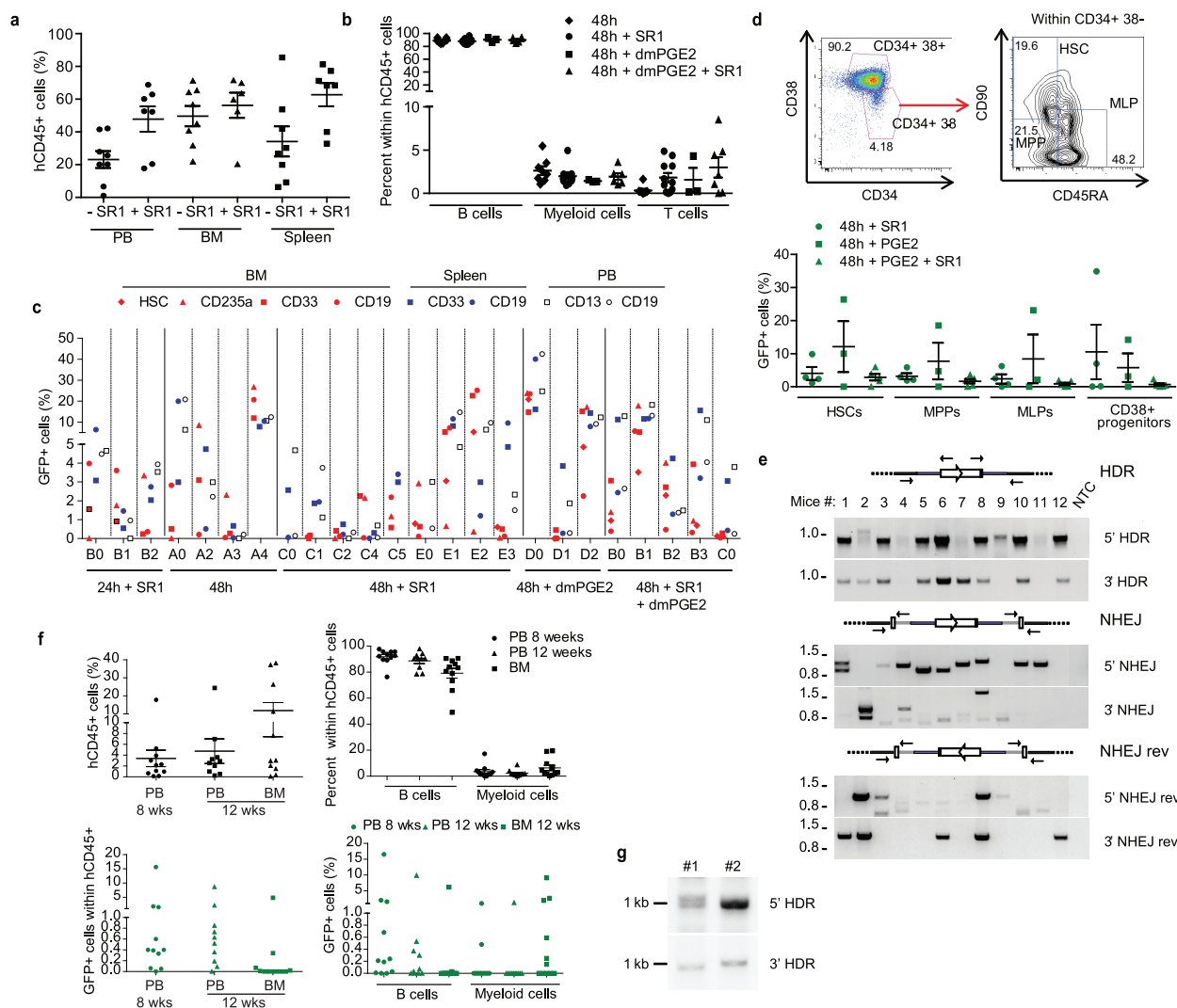
Means  $\pm$  s.e.m.;  $n = 2$  or 3, respectively. Overall, there was a transient reduction in viable cell number 24 h after electroporation, also observed in CFC yield, which resulted from the combined exposure to electroporation, ZFN mRNA and IDLV. However, the surviving cells grew with similar kinetics as the untreated controls in liquid culture and gave rise to similar proportions of myeloid and erythroid colonies. **f, g**, Targeting specificity of integration. **f**, Southern blot (top) and PCR (bottom) analyses for targeted integration into *IL2RG* on iPSCs obtained by reprogramming GFP<sup>+</sup> cells from Fig. 1c. UT, untreated cells. **g**, Genomic DNA from GFP<sup>+</sup> colonies was analysed by PCR for targeted integration into *AAVS1* or *IL2RG*. The gels show the PCR amplicons for either the 5' or 3' HDR integration junction at the genomic target site and a control locus (*CCR5*). The percentages of colonies positive for targeted integration by PCR are reported in Fig. 1d.



### Extended Data Figure 3 | Investigating lower gene-targeting efficiency in the more primitive cells.

**a**, Gating strategy used to identify subpopulations of cord blood cells according to expression of CD90, CD133 and CD34 surface markers. **b**, After 24 h of pre-stimulation, CD34<sup>+</sup> cells were electroporated with 175 µg of GFP mRNA (the selected dose of a ZFN mRNA from Extended Data Fig. 1). Flow cytometry analysis was performed 2 days later using the gating strategy shown in Fig. 3a. Bars represent the percentage of GFP<sup>+</sup> cells (plotted on left axis) while the line shows the level of transgene expression (plotted on the right axis as MFI, measured in arbitrary units). Means  $\pm$  s.e.m. ( $n = 16$  on 6 cord blood donors). **c**, CD34<sup>+</sup> cells treated for targeted integration were sorted by FACS one day after electroporation according to the gating strategy shown in Fig. 3a. The sorted populations were sampled at the indicated times and levels of NHEJ at the ZFN target site (AAVS1) were determined by Cell

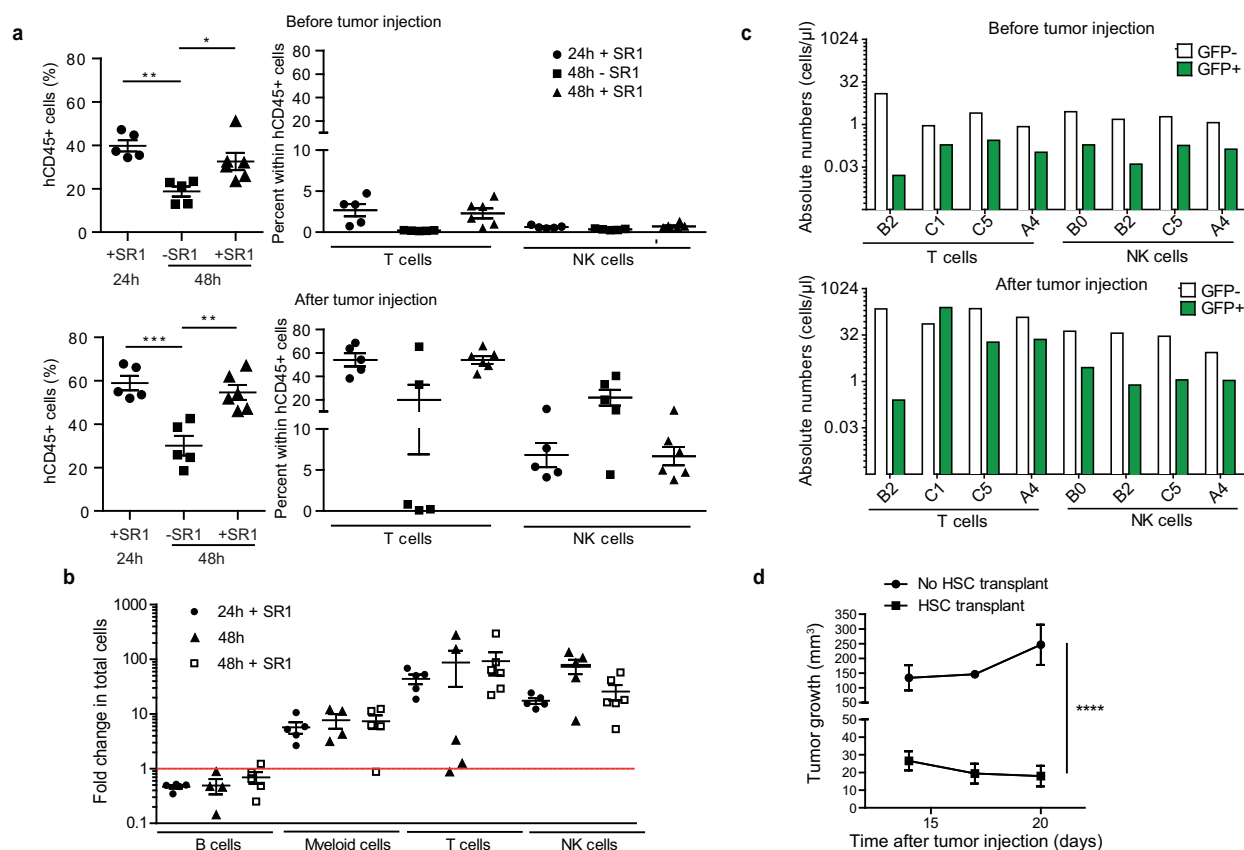
assay ( $n = 3$ ). **d**, Apoptosis analysis performed one day after electroporation on CD34<sup>+</sup> cells transduced with IDLV and electroporated with ZFN mRNAs. Percentages of live (7AAD<sup>-</sup>, AnnexinV<sup>-</sup>), early apoptotic (7AAD<sup>-</sup>, AnnexinV<sup>+</sup>), late apoptotic (7AAD<sup>+</sup>, AnnexinV<sup>+</sup>) and necrotic (7AAD<sup>+</sup>, AnnexinV<sup>-</sup>) cells. Means  $\pm$  s.e.m. ( $n = 5$  on 4 cord blood donors). **e**, Representative growth curves of CD34<sup>+</sup> cells cultured in the presence or not of SR1 and treated (TI) or not (UT) with the protocol described in Fig. 1b. Note that the TI and UT growth curves are reproduced from Extended Data Fig. 1c. Addition of SR1 to the culture did not change the proliferation rate of the cells. Because SR1 did not increase the total number of cells but the percentage of more primitive cells in culture (Fig. 3c), the absolute number of primitive cells is larger in cultures containing SR1.



**Extended Data Figure 4 | Long-term multilineage engraftment of gene-targeted CD34<sup>+</sup> cells in primary and secondary NSG mice.** **a**, Percentages of human cells in the indicated organs of NSG mice 15–23 weeks after transplantation with CD34<sup>+</sup> cells treated with the improved protocols from Fig. 3a with or without SR1. **b**, Percentage of the indicated lineages within human cells in peripheral blood of mice 14 weeks post-transplant. Means  $\pm$  s.e.m. (48 h,  $n = 8$ ; 48 h SR1,  $n = 11$ ; 48 h PGE2,  $n = 3$ ; 48 h PGE2 SR1,  $n = 6$ ). Overall, the addition of SR1 and PGE2 to the *in vitro* culture did not significantly affect the *in vivo* differentiation of treated cells. Notably, the increased human engraftment achieved with the optimized culture conditions (as illustrated in Fig. 3g) correlates with increased T-cell output. **c**, Multilineage GFP marking in individual NSG mice transplanted with CD34<sup>+</sup> cells treated with the indicated protocols for targeted integration. Percentages of GFP<sup>+</sup> cells were calculated within the CD45<sup>+</sup> Lin<sup>+</sup> populations (represented with different data-point shapes) in different organs (represented by different data-point colours). The analysis was performed on peripheral blood 14 weeks after transplantation and on spleen and bone marrow at the end of the experiments. Only mice displaying  $\geq 0.1\%$  GFP<sup>+</sup> cells are represented in the graph ( $n = 2$  independent experiments). Note that when using the improved protocols for targeted integration GFP<sup>+</sup> cells are found in multiple lineages in all mice. **d**, Analysis of the primitive human compartment in the bone marrow of transplanted mice from **c**. Top: gating strategy used to define progenitors (CD34<sup>+</sup> CD38<sup>+</sup>), multilymphoid progenitors (MLPs; CD34<sup>+</sup> CD38<sup>+</sup> CD90<sup>low</sup> CD45RA<sup>+</sup>), multipotent progenitors (MPPs; CD34<sup>+</sup> CD38<sup>+</sup> CD90<sup>low</sup> CD45RA<sup>+</sup>) and HSCs (CD34<sup>+</sup> CD38<sup>+</sup> CD90<sup>low</sup> CD45RA<sup>+</sup>). Bottom: percentages of GFP<sup>+</sup> cells within the defined populations. Means  $\pm$  s.e.m. (48 h SR1,  $n = 4$ ; 48 h PGE2,  $n = 3$ ; 48 h PGE2 SR1,  $n = 5$ ). **e**, DNA from total bone marrow cells of transplanted mice was analysed by PCR to determine TI into *IL2RG*. Each column represents one mouse. The schematics show the different sets of primers used to detect on-target insertions mediated by HDR or NHEJ

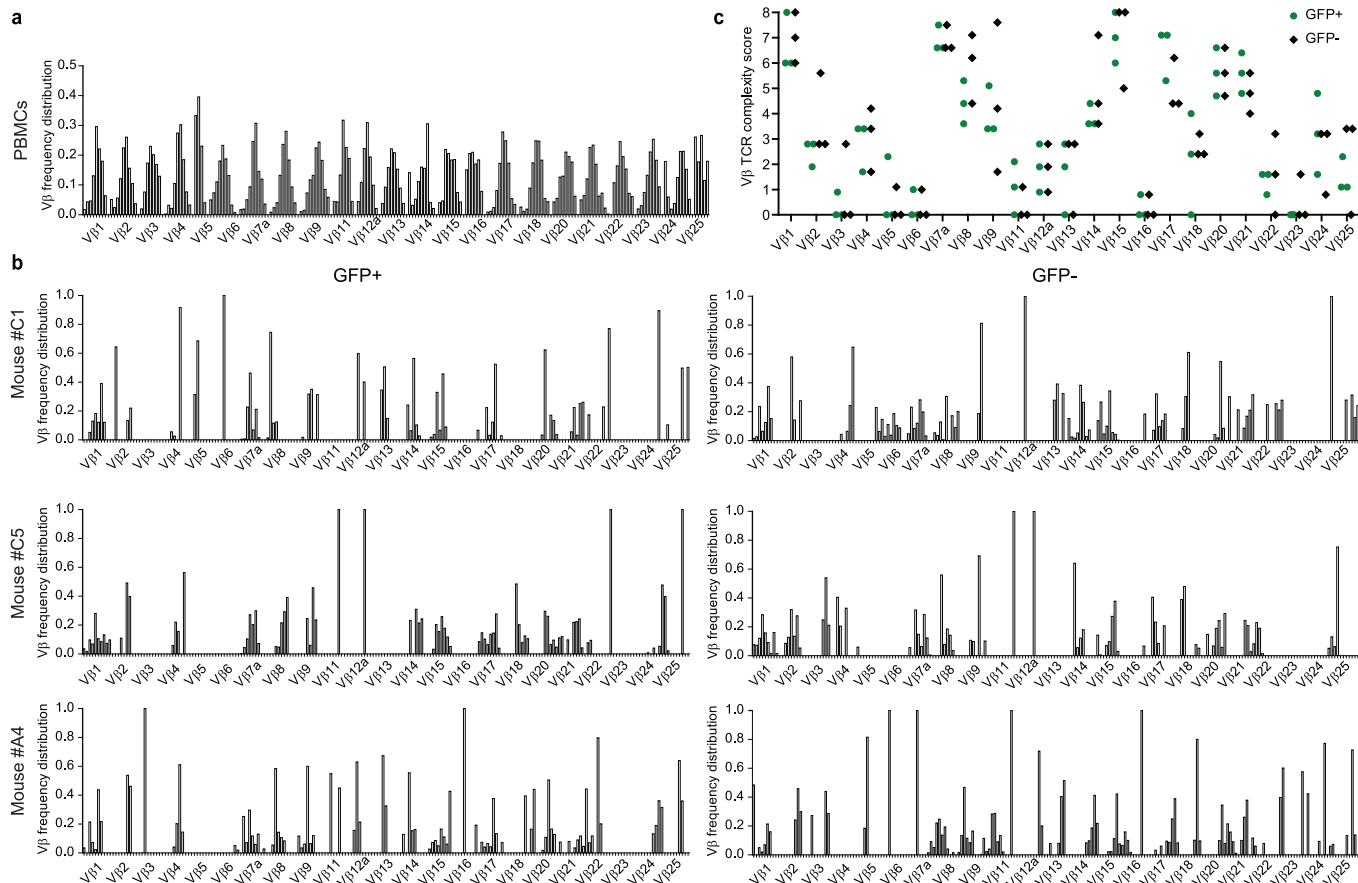
(the latter shown with the vector in sense or reverse orientation with respect to *IL2RG*). Whereas evidence of HDR-mediated insertion of the cassette was retrieved from all mice assayed, there was also indication of some NHEJ-mediated integration of the donor IDLV. Trapping of IDLV at sites of NHEJ has been previously reported<sup>16</sup>. We note that our strategy could be adjusted to also exploit this type of insertion to drive transgene expression<sup>24</sup> and potentially increase the overall efficiency of gene correction (see schematic in Fig. 4a). In such case, one should target insertion within an intron of the gene so that the splice acceptor site of the corrective cDNA is next in line for processing with the splice donor site of the upstream endogenous exon and any intervening sequence can be spliced out from the chimaeric transcript leading to reconstitution of a functional open reading frame (provided that the insertion occurred in the same orientation as gene transcription). An additional benefit of targeting an intronic sequence would be to spare exons from disruption by NHEJ, although this is of minor concern when dealing with already defective alleles. **f**, 15–23 weeks after the primary transplant, human CD34<sup>+</sup> cells were purified from the bone marrow of 11 mice from **c** (mice were chosen among those best engrafted with GFP<sup>+</sup> cells in the different CD34<sup>+</sup> cell treatment groups) and transplanted (one mouse to one mouse) into 7–11-week-old NSG mice. Secondary recipient mice were monitored for engraftment of human CD45<sup>+</sup> and GFP<sup>+</sup> cells at 8 and 12 weeks post-transplant in peripheral blood, and on bone marrow at the end of the experiments. Top left: percentages of human cells in the indicated organs. Top right: percentages of the indicated lineages within the human cells. B cells were defined by expression of CD19 and myeloid cells were defined by expression of CD13 in peripheral blood or CD33 in bone marrow. Dots represent individual mice. Bottom left: percentages of GFP<sup>+</sup> cells within the human graft in the indicated organs and (bottom right) within the indicated lineages, as on top. **g**, Human cells from the bone marrow of the top 2 engrafted secondary recipient mice from **f** were analysed by PCR for targeted integration into *IL2RG*.





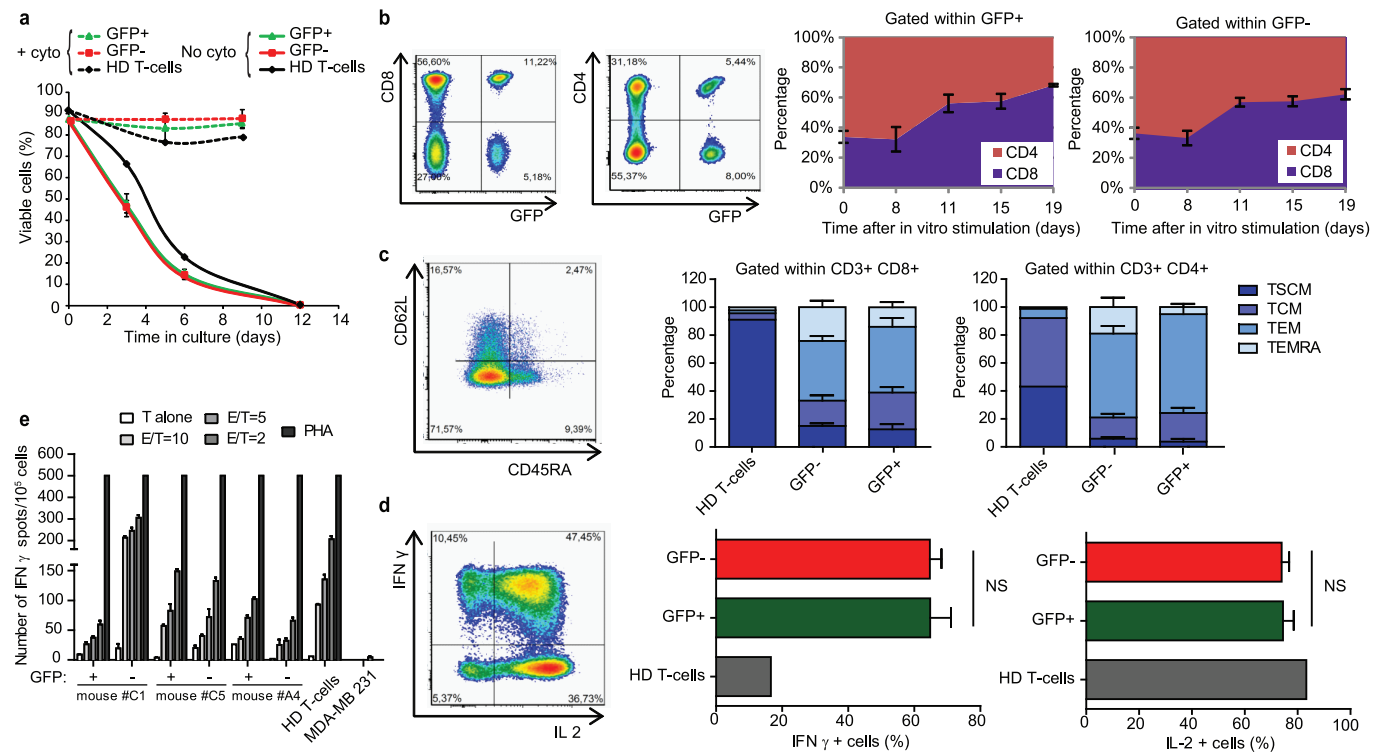
**Extended Data Figure 5 | Expansion of lymphoid cells in the transplanted mice after tumour challenge.** **a**, Left: percentage of human cells in the peripheral blood of mice transplanted with male CD34<sup>+</sup> cells treated as indicated for targeted integration into *IL2RG*. Analyses performed at the time of tumour injection (top panel) and 3 weeks later (bottom panel). Right: percentages of T and NK cells (CD3<sup>+</sup> and CD16/56<sup>+</sup> cells, respectively) measured within the human CD45<sup>+</sup> cells in peripheral blood. **b**, Fold change in the absolute number of the indicated lineages in peripheral blood 3 weeks after

tumour challenge. Means  $\pm$  s.e.m. (24 h SR1,  $n = 5$ ; 48 h SR1,  $n = 6$ ; 48 h,  $n = 5$ ). **c**, Counts of GFP<sup>+</sup> and GFP<sup>-</sup> T and NK cells in the peripheral blood of transplanted NSG mice before (top) and 3 weeks after (bottom) injection of the MDA-MB 231 tumour cell line engineered to express human IL-7, IL-15 and GM-CSF. Fold changes calculated from these values are plotted in Fig. 4e. **d**, Tumour growth in mice transplanted ( $n = 16$ ) or not ( $n = 3$ ) with treated CD34<sup>+</sup> cells. \*\*\*\* $P < 0.0001$  (two-way ANOVA).



**Extended Data Figure 6 | Generation of T cells with substantial Vβ TCR diversity from the engrafted gene-targeted HSPCs.** **a**, Analysis of TCR Vβ repertoire performed on PBMCs from a healthy donor. The histogram shows the frequency distribution of the different complementarity-determining region 3 (CDR3) lengths identified within the indicated Vβ families. As expected from a highly polyclonal TCR repertoire, all Vβ families display a Gaussian distribution of the CDR3 lengths. **b**, Frequency distribution of the different CDR3 regions measured as in **a** for the GFP<sup>+</sup> (left) and GFP<sup>-</sup> (right) T cells harvested from the transplanted C1, C5 and A4 mice from Extended Data

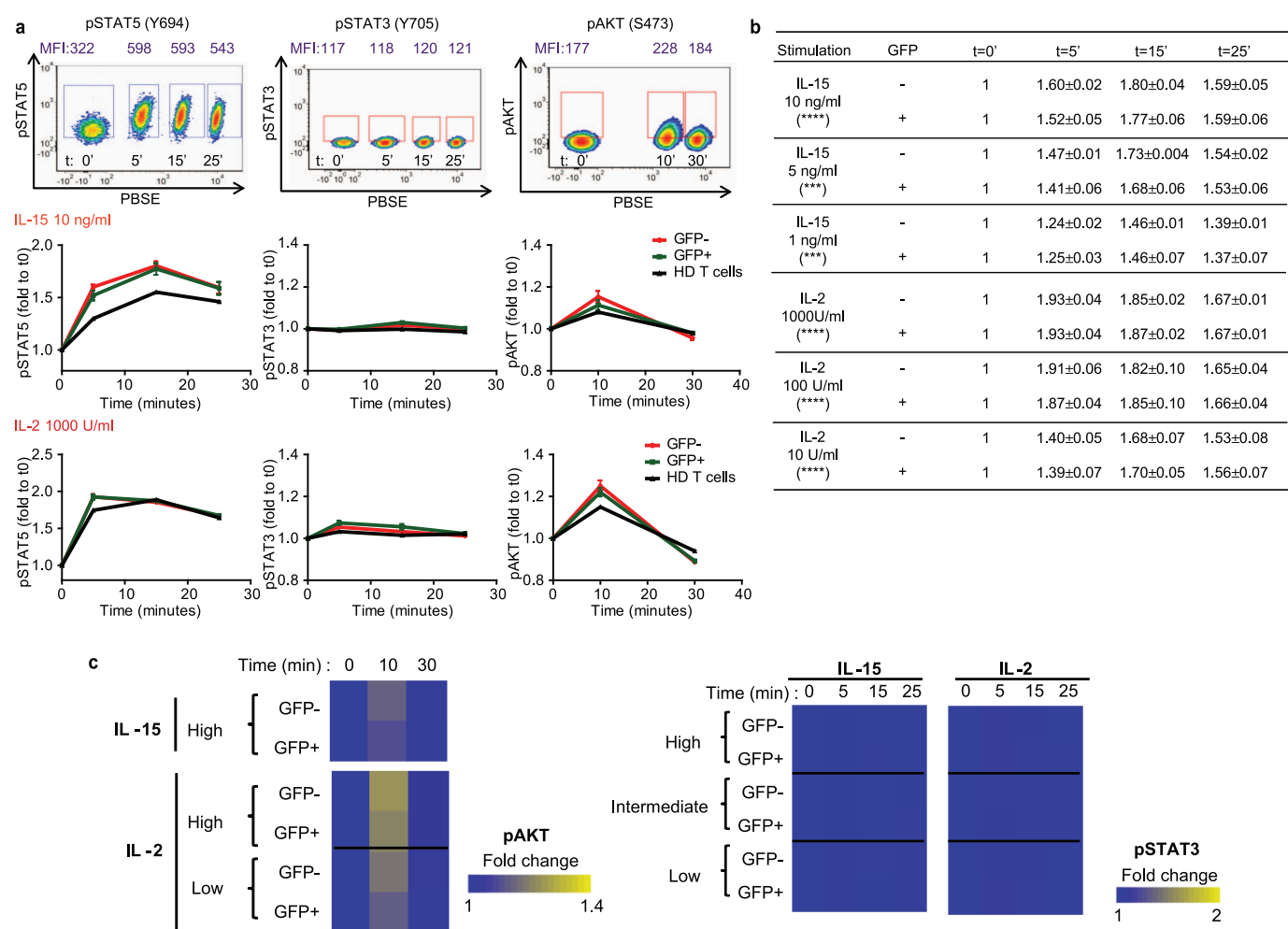
Fig. 5c. **c**, Complexity score<sup>36</sup> assigned to each Vβ family for the samples shown in **b**. The sum of the scores for all the family is plotted in Fig. 4g. Note that all the samples analysed display similar TCR Vβ repertoire distributions, constrained for some families and more polyclonal for others, as might be expected for human T cells developed in haematochimaeric mice<sup>41</sup>, and that no significant differences are observed between the GFP<sup>+</sup> and GFP<sup>-</sup> cells. This finding indicates that the rate of gene targeting achieved in the transplanted stem/progenitor cells does not detectably limit the generation of a polyclonal and functional T-cell repertoire *in vivo*.



**Extended Data Figure 7 | Functional and phenotypic characterization of *IL2RG*-edited T cells harvested from transplanted NSG mice.** **a**, Graph showing the viability of GFP<sup>+</sup> and GFP<sup>-</sup> T cells harvested from the transplanted NSG mice (from Extended Data Fig. 6) or of T cells from peripheral blood of healthy donor (HD T cells), cultured in the presence (+ cyto) or absence (no cyto) of human IL-7 and IL-15. **b**, Left: representative density plots of GFP<sup>+</sup> and GFP<sup>-</sup> T cells harvested from mice, stained for CD8 (left) and CD4 (right). Right: GFP<sup>+</sup> and GFP<sup>-</sup> T cells harvested from mice were activated *ex vivo* with beads coated with anti-CD3 and anti-CD28-specific antibodies, and cultured with IL-7 and IL-15. CD4 and CD8 composition of GFP<sup>+</sup> and GFP<sup>-</sup> cells, measured during *ex vivo* culture, is shown ( $n = 3$ ). **c**, Surface phenotype of CD4 and CD8 T cells from **b** and HD T cells at day 19

after stimulation. A representative plot (left) and histograms with medians + s.e.m. (right) are shown. T stem memory cells (TSCM) are defined as CD62L<sup>+</sup> CD45RA<sup>+</sup> (refs 42, 43), T central memory (TCM) as CD62L<sup>+</sup> CD45RA<sup>-</sup>, T effector memory (TEM) as CD62L<sup>-</sup> CD45RA<sup>-</sup> and terminal effectors (TEMRA) as CD62L<sup>-</sup> CD45RA<sup>+</sup>. **d**, Production of IL-2 and IFN- $\gamma$  by GFP<sup>+</sup> or GFP<sup>-</sup> T cells as in **b**, and by HD T cells after 6 h stimulation with PMA+ionomycin. A representative plot (left) and percentages of IL-2<sup>+</sup> and IFN- $\gamma$ <sup>+</sup> cells are shown.  $P = ns$  (unpaired *t*-test). **e**, IFN- $\gamma$  Elispot assay showing the frequencies of IFN- $\gamma$ -producing cells from **b** challenged with the MDA-MB 231 tumour cell line at different effector to target ratios. PHA stimulation was used as positive control.





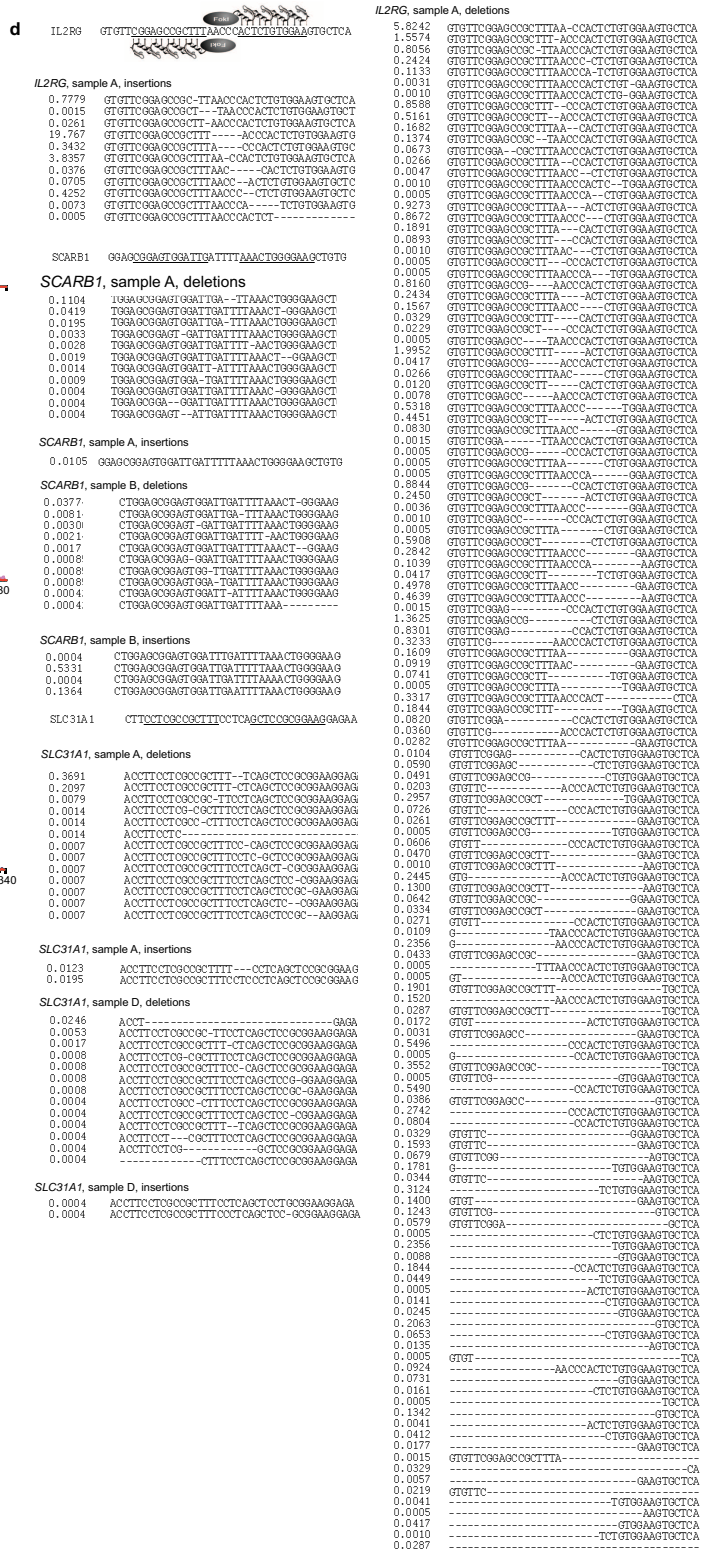
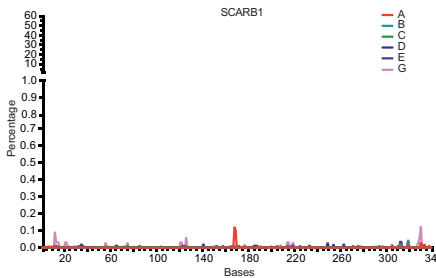
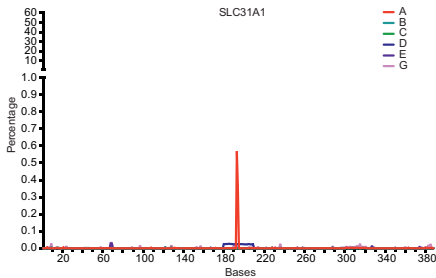
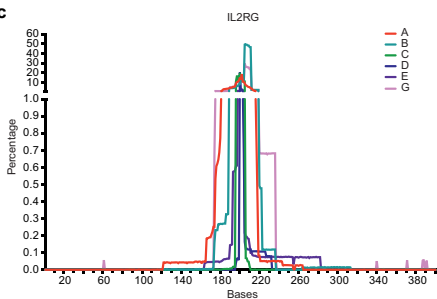
**Extended Data Figure 8 | Functionality of  $\gamma$ -chain-dependent signalling pathway in *IL2RG*-gene-edited T cells.** GFP<sup>+</sup> or GFP<sup>-</sup> T cells from the transplanted mice (as in Extended Data Figs 6 and 7) or T cells from the peripheral blood of healthy donor (HD T-cells) were exposed to the indicated doses of  $\gamma$ -chain-related cytokines. The phosphorylation levels of STAT5 on Y694 (pSTAT5), STAT3 on Y705 (pSTAT3) and AKT on S473 (pAKT) were measured at the indicated time points by flow cytometry analyses. **a**, Top: representative plots showing pSTAT5, pSTAT3 and pAKT. Each time point of analysis is labelled by a different concentration of the intracellular dye PBSE.

Bottom: fold changes in the levels of pSTAT5, pSTAT3 and pAKT relative to the time 0 of stimulation with the indicated maximal doses of IL-15 and IL-2 ( $n = 3$ ). Phosphorylation of STAT3 was used as a specificity control.  $P = ns$  (two-way ANOVA). **b**, Table showing the fold changes in the levels of pSTAT5 which are graphically represented in Fig. 4k. Statistics in the first column indicate significant change during time. \*\*\*\* $P < 0.0001$ , \*\*\* $P < 0.001$  (two-way ANOVA). **c**, Heat map representing fold changes in pAKT and pSTAT3 levels after the indicated time of exposure (min) to increasing amounts of IL-2 or IL-15.  $P = ns$  between GFP<sup>+</sup> and GFP<sup>-</sup> cells (two-way ANOVA).

a	% Indels	In Vitro			Mouse			UT
		A	B	G	C	D	E	
Region								
IL2RG	54.843%	61.419%	45.843%	26.321%	43.745%	20.310%	0.240%	
SCARB1	0.194%	0.727%	0.031%	0.038%	0.034%	0.060%	0.028%	
SLC31A1	0.628%	0.016%	0.057%	0.018%	0.039%	0.011%	0.017%	
FAM133B	0.113%	0.000%	0.042%	0.062%	0.047%	0.000%	0.054%	
KIAA0528	0.022%	0.089%	0.000%	0.027%	0.020%	0.024%	0.027%	
SF3B1	0.029%	0.035%	0.033%	0.025%	0.042%	0.039%	0.028%	
A2BP1	0.012%	0.000%	0.006%	0.012%	0.012%	0.011%	0.005%	
ANKFY1	0.022%	0.018%	0.020%	0.019%	0.022%	0.018%	0.017%	
TRIM43	0.005%	0.006%	0.000%	0.009%	0.006%	0.004%	0.006%	
SEC16A	0.020%	0.016%	0.000%	0.000%	0.017%	0.000%	0.041%	

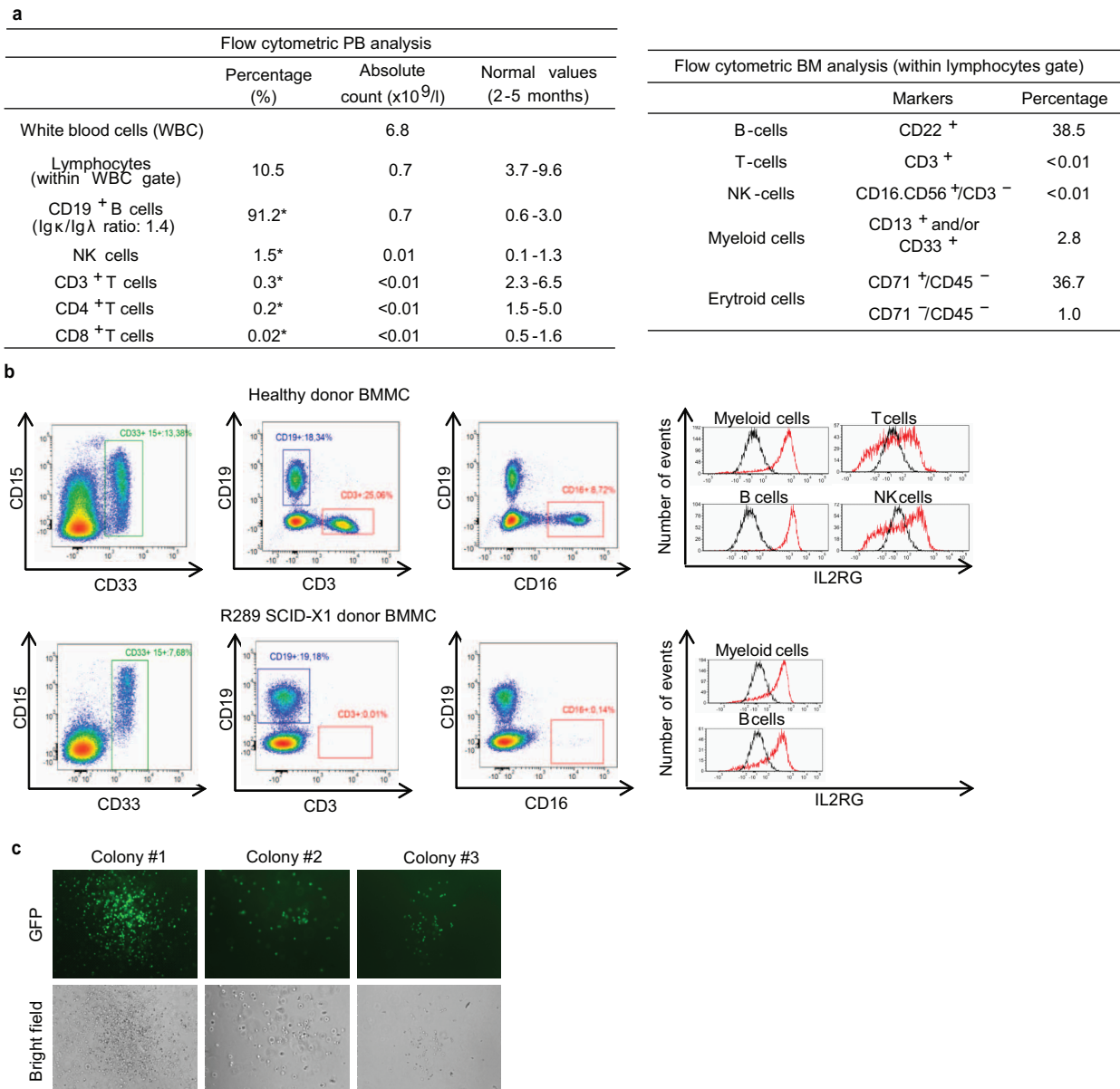
b	P-value	In Vitro			Mouse		
		A	B	G	C	D	E
Region							
IL2RG	1.00E-10	1.00E-10	1.00E-10	1.00E-10	1.00E-10	1.00E-10	1.00E-10
SCARB1	0.00025	2.44E-10	—	0.1154	1	—	0.49721
SLC31A1	2.10E-10	—	0.79512	0.0055	1	5.41E-05	0.15933
FAM133B	0.17656	0	0.06697	0.69641	—	1	1
KIAA0528	0.31701	0.07256	1	0.92092	0.09423	0.59396	0.59396
SF3B1	0.77826	0.23094	0.41487	0.7472	0.02115	0.06942	0.06942
A2BP1	0.05599	0.33634	0.82108	0.04689	0.08366	0.09301	0.09301
ANKFY1	0.43578	1	0.66383	0.79064	0.36158	0.88084	0.88084
TRIM43	0.62967	1	1	0.40318	1	0.47769	0.47769
SEC16A	0.39374	0.48659	—	1	0.49411	1	1



**Extended Data Figure 9 | Indel quantification on the intended ‘on’ target and candidate ‘off’ target genomic sites of *IL2RG* ZFNs.** **a**, Indel quantification expressed as percentage of the total number of reads for each sample and target. The untreated (UT) sample F is shown in the rightmost column. 0 values underlie undetected events due to low sequencing read coverage (see Supplementary Information). **b**, Table of *P*-values obtained comparing the number of indels for each target and its untreated control sample (Fisher’s exact test for contingency data). *P*-values allowed distinguishing real accumulation of indels (labelled with a green arrow close to the *P*-value) from background noise (marked with a yellow horizontal bar). Only *SCARB1* and *SLC31A1* show indels in samples A, B and D at a frequency significantly higher than the untreated sample. **c**, Indel frequency distribution along the amplified genomic sequences of the positive control *IL2RG* and the

two loci that showed a low but significant off target activity, *SCARB1* and *SLC31A1*. The *x* axis represents the amplified sequence in base pairs while the *y* axis shows for each base the percentage of reads that reported indels for each sample after noise subtraction (see Supplementary Information). Note that indels mainly occur in the central region of the amplicon, corresponding to the spacer between the genomic sequences expected to be bound by the ZFNs. **d**, Representative sequence alignments of retrieved indels in *IL2RG* (sample A), *SCARB1* (from samples A and B) and *SLC31A1* (from samples A and D) focusing the analysis on the region bearing identity or homology to the intended ZFN target sequence. For each sequence type the relative frequency of retrieval in the sample is reported as a percentage. As shown in **c** indels mostly occur in the central spacer region between the 2 ZFN binding sites (underlined).





**Extended Data Figure 10 | *IL2RG* gene correction in CD34<sup>+</sup> cells from the bone marrow of a genotyped subject with SCID-X1.** **a**, Blood cell counts in the peripheral blood and bone marrow of a SCID-X1 male child carrying the R289X mutation in the *IL2RG* gene (see Supplementary Information) and showing virtual absence of T and NK cells. Asterisks in the left-most table indicate values calculated within the leukocyte gate. **b**, Left: representative density plots showing the cellular composition of a bone marrow harvest from a healthy donor (top) and the subject with SCID-X1 (bottom) after purification of nucleated cells (bone-marrow-derived mononuclear cells, BMMCs). Myeloid cells are stained for CD15 and CD33, B cells for CD19, T cells for CD3 and NK cells for CD16. Right:  $\gamma$ -chain expression within the indicated cell

populations gated from the plots shown on the left. As expected from the missense R289X mutation, the  $\gamma$ -chain protein is expressed on the cell surface but it is not functional, as indicated by the absence of T and NK cells in the patient. In Fig. 5 we show normal expression of the  $\gamma$ -chain protein in the myeloid cell progeny of three gene-corrected CFCs from the patient cells (gene correction was proven by evidence of targeted integration in the only *IL2RG* allele of this male individual in the clonal CFC progeny and by the expression of the fusion transcript bearing the corrective cDNA). These data indicate normal  $\gamma$ -chain expression from the reconstituted allele in the edited patient cells. **c**, Bright-field and fluorescence microscopy images of the three GFP<sup>+</sup> myeloid colonies obtained after *IL2RG* gene correction of SCID-X1 cells.

# Homologue engagement controls meiotic DNA break number and distribution

Drew Thacker<sup>1,2</sup>, Neeman Mohibullah<sup>1,3</sup>, Xuan Zhu<sup>1,2</sup> & Scott Keeney<sup>1,2,3</sup>

Meiotic recombination promotes genetic diversification as well as pairing and segregation of homologous chromosomes, but the double-strand breaks (DSBs) that initiate recombination are dangerous lesions that can cause mutation or meiotic failure. How cells control DSBs to balance between beneficial and deleterious outcomes is not well understood. Here we test the hypothesis that DSB control involves a network of intersecting negative regulatory circuits. Using multiple complementary methods, we show that DSBs form in greater numbers in *Saccharomyces cerevisiae* cells lacking ZMM proteins, a suite of recombination-promoting factors traditionally regarded as acting strictly downstream of DSB formation. ZMM-dependent DSB control is genetically distinct from a pathway tying break formation to meiotic progression through the Ndt80 transcription factor. These counterintuitive findings suggest that homologous chromosomes that have successfully engaged one another stop making breaks. Genome-wide DSB maps uncover distinct responses by different subchromosomal domains to the ZMM mutation *zip3* (also known as *cst9*), and show that Zip3 is required for the previously unexplained tendency of DSB density to vary with chromosome size. Thus, feedback tied to ZMM function contributes in unexpected ways to spatial patterning of recombination.

DSBs are hazardous genomic damage that most cells avoid but that each meiotic cell introduces in large numbers; therefore the activity of Spo11, the protein that makes DSBs, must be tightly regulated<sup>1</sup>. Typical depictions of recombination pathway(s) (Fig. 1a) implicitly divide involved proteins into upstream (DSB formation) and downstream (DSB repair) factors. This view suggests that eliminating downstream factors will have little or no effect on number or distribution of upstream events (DSBs). An alternative view considers recombination genome-wide, not just at any one site: DSBs do not form all at once, so fates of early DSBs might govern whether and where later DSBs form. In this scenario, downstream factors may behave genetically as upstream factors if their absence disrupts feedback circuits.

Precedents for feedback are known in several organisms. In mice, yeast and flies, ATM kinase governs a negative feedback loop inhibiting DSB formation in response to breaks<sup>2–5</sup>. In mice, flies and worms, defective interhomologue interactions are known or proposed to allow continued DSB formation, suggesting another type of feedback<sup>6–8</sup>. The logic of these circuits predicts different behaviour in DSB repair mutants: the ATM-type circuit should suppress further DSB formation if existing breaks cannot be repaired, but, conversely, defective interhomologue interactions caused by repair defects might instead allow more DSBs to accumulate. We test these predictions here and ask whether feedback contributes to the spatial organization of recombination. We focus on the ZMM proteins in *Saccharomyces cerevisiae* (Zip1–4 (Zip4 also known as Spo22), Msh4–5, Mer3, Spo16 and Pph3). These biochemically diverse factors shepherd recombination intermediates towards a crossover fate and help build synaptonemal complexes, therefore ZMM null mutations cause recombination and synaptonemal complex defects, with varying degrees of meiotic arrest<sup>9</sup>.

## Increased DSB numbers in ZMM mutants

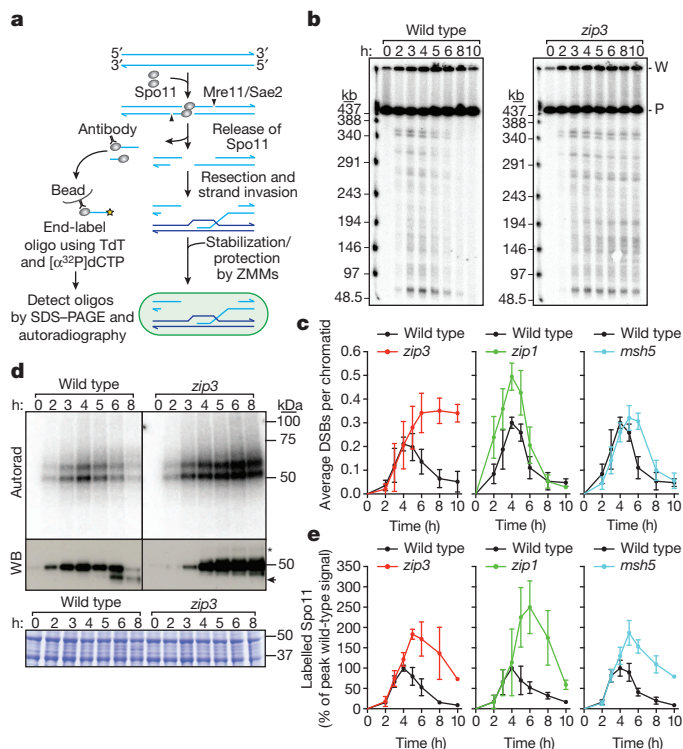
We measured DSBs by Southern blotting whole chromosomes separated on pulsed-field gels (Fig. 1b, c). In wild type, chromosome fragments appeared and disappeared as DSBs were formed and repaired, as expected.

By contrast, broken chromosomes accumulated in *zip3* mutants for at least 2 h after DSBs waned in wild type, reaching a plateau 1.7-fold higher than the wild-type peak and persisting for hours (Fig. 1c). (This underestimates DSBs because *zip3* mutants complete some repair (below).) In *zip1* mutants, broken chromosomes reached higher levels than wild type before disappearing, and in *msh5* mutants, time-averaged DSB levels were higher than in wild type but the maximum was only slightly increased (Fig. 1c and Extended Data Fig. 1). Differences in arrest may account for variation between mutants at later times (Supplementary Discussion). In principle, increased steady-state DSBs could reflect extended lifespan, increased frequency, or both. As these measurements cannot distinguish between these possibilities, we applied a battery of methods that offset limitations of any one approach.

To assess DSBs globally and mitigate uncertainty from repair defects, we examined Spo11–oligonucleotide (oligo) complexes, by-products of DSB formation that can be used to measure DSB number and distribution<sup>2,10,11</sup> (Fig. 1a). Extracts were prepared from cultures expressing phenotypically normal Flag-tagged Spo11 (Extended Data Fig. 2). Anti-Flag immunoprecipitates were labelled with terminal transferase and [ $\alpha$ -<sup>32</sup>P]dCTP, resolved by SDS–polyacrylamide gel electrophoresis (SDS–PAGE), then labelled Spo11–oligo complexes were detected by phosphorimager and total Spo11 was detected by western blotting (Fig. 1d). In wild type, Spo11–oligo complexes appeared contemporaneously with DSBs, peaked at ~4 h, then declined (Fig. 1d, e). In *zip3* mutants, Spo11–oligo complexes first appeared with similar timing and levels as wild type, but continued to accumulate after 4 h (Fig. 1d, e). Spo11–oligo levels reached a maximum at ~5 h that was 1.8-fold higher than the wild-type peak and remained high after most complexes had disappeared in wild type (altered accumulation of free Spo11 protein is probably due to arrest (Supplementary Discussion)). Similar patterns were seen in *msh5* and *zip1* mutants (Fig. 1e and Extended Data Fig. 3). If turnover of DSBs and Spo11–oligo complexes is separable, these findings imply that ZMM mutants make more breaks.

<sup>1</sup>Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>2</sup>Weill Graduate School of Medical Sciences of Cornell University, New York, New York 10065, USA.

<sup>3</sup>Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA.



**Figure 1 | More DSBs form in ZMM mutants.** **a**, Spo11 generates a covalent protein-linked DSB; endonucleolytic cleavage releases Spo11 bound to a short oligo (detection method on left). Resection is followed by strand invasion and ZMM-dependent stabilization of intermediates fated to become crossovers. **b**, **c**, Representative pulsed-field gel Southern blots probed for chromosome IX are shown in **b** and Poisson-corrected DSB quantification shown in **c** (mean  $\pm$  s.d., 3 cultures). P, parental; W, wells. **d**, **e**, Representative Spo11-oligo complex time courses are in **d** and quantification in **e** (mean  $\pm$  s.d. for 3 cultures, except at 10 h for *msh5* and *zip3* analyses (1 culture)). Radiolabelled Spo11-oligo complexes were detected by autoradiography (top) and total Spo11 was detected by anti-Flag western blot (WB, middle). The main labelled species differ in oligo size<sup>10</sup>. Nearly all of the western blot signal is Spo11 that has not made a DSB<sup>10</sup>. Asterisk indicates species co-migrating with upper Spo11-oligo complexes; arrowhead represents proteolytic product. Bottom, extract samples run separately and stained with Coomassie control for input to the immunoprecipitates. In panels **c** and **e**, mutants are plotted with wild-type data collected in parallel.

If ZMM mutants make more breaks, then more repair products should also accumulate. To gauge interhomologue recombination, we used strains heterozygous for different *arg4* mutations (Fig. 2a). Prophase cells transferred to rich medium abort meiosis, often completing recombination even if unable while still in meiosis<sup>12</sup>. All ZMM mutants tested except *msh5* formed more Arg<sup>+</sup> prototrophs than wild type (Fig. 2b). Increased recombination has been reported in all ZMM mutants examined (including *msh4* and *msh5*), but was not interpreted as evidence for increased DSB frequency<sup>13–19</sup> (Supplementary Table 1). Thus, a context-dependent hyper-rec phenotype is a common but previously unrecognized property of ZMM mutants.

We explored this hyper-rec behaviour by quantifying recombinants at three natural DSB hotspots (Fig. 2c and Extended Data Fig. 4a, b). At each, allelic copies have different flanking and central restriction sites. Crossovers and parental length fragments are resolved by electrophoresis after digestion with flanking enzymes, then DNA is digested in the gel with the central enzyme before electrophoresis in the orthogonal dimension. Noncrossover gene conversion molecules co-migrate with one parent in the first dimension but have a central restriction site matching the other parent (Fig. 2d and Extended Data Fig. 4c–e). Key features are that the hotspots are high-intensity with few/no other DSB sites nearby, and central polymorphisms are positioned to make incorporation into

heteroduplex DNA likely. At *CCT6* and *ERG1*, recombinant molecules were 1.7–2.5-fold more abundant in the ZMM mutants tested, with increased noncrossovers and crossovers at or below wild-type levels (Fig. 2e). At *GAT1*, *zip3* mutants displayed fewer crossovers offset by more noncrossovers, for a net frequency comparable to wild type (Fig. 2e). (Refer to Supplementary Discussion of gene conversion tracts and sister chromatid recombination.) These findings reinforce the conclusion that ZMM mutations cause a hyper-rec phenotype that is variable between loci.

DSBs were present at *CCT6* and *GAT1* in *zip3* mutants at late times, well past the point when most DSBs had disappeared in wild type (Fig. 2f, g and Extended Data Fig. 5a). Recombination intermediates ('joint molecules') that are transient in wild type were also detected late in *zip3*. DSBs and joint molecules at *CCT6* were also present in *zip1* and *msh5* mutants later than in wild type but not as long as in *zip3*, similar to analysis of breaks on chromosome IX (Extended Data Fig. 5b–e). These results agree with data at artificial hotspots in ZMM mutants (for example, ref. 20), but it was not previously possible to evaluate whether DSB numbers were increased and most previous studies dismissed or did not consider this possibility (Supplementary Table 1). We can now combine DSB data with quantification of recombination intermediates and products (Supplementary Table 2): this bookkeeping reveals that *msh5*, *zip1* and *zip3* mutants experience 1.8–2.6-fold more detectable DSB-related events at *CCT6*. Recombination product overabundance yields the same conclusion for *zip3* and *msh5* at *ERG1* (1.7–1.9-fold), whereas wild type and *zip3* mutants had similar totals at *GAT1*. We conclude that ZMM mutants incur more DSBs, but to varying degrees at different loci.

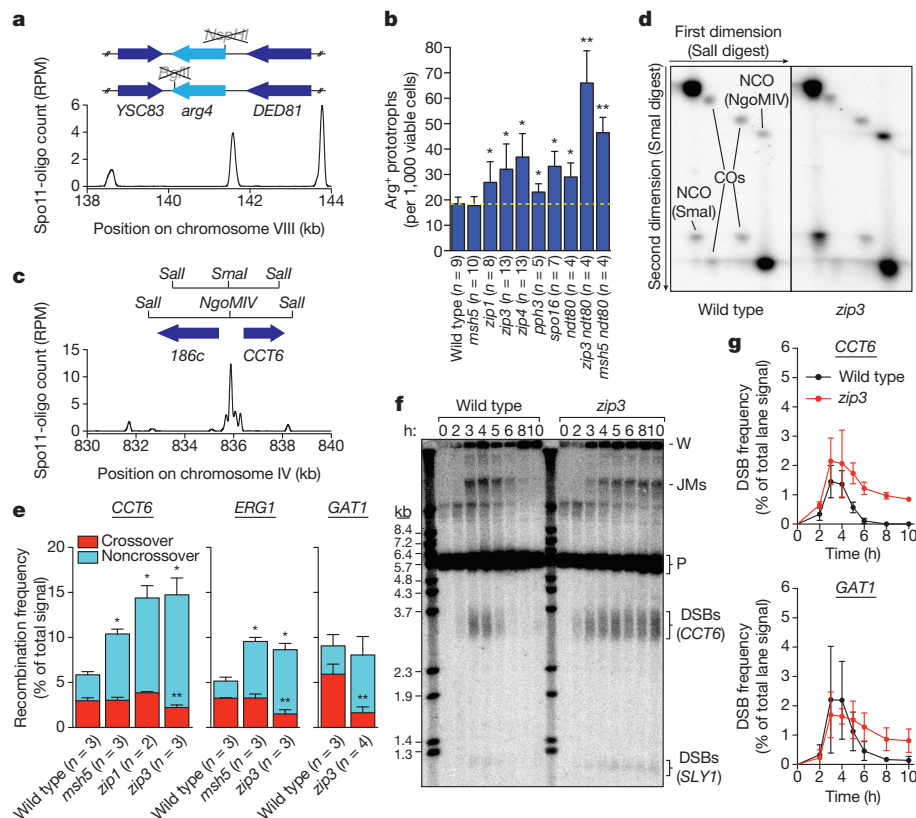
### Separate pathways controlling DSB number

Recombination products and DSBs accumulate in cells lacking Ndt80, a transcription factor controlling pachytene exit<sup>12,21</sup>, and therefore it has been suggested that this stage in prophase ends a period permissive for DSB formation<sup>1,21</sup>, further supported by recent studies<sup>5,22,23</sup>. Indeed, Spo11-oligo complexes reached 1.2–1.4-fold higher than the wild-type maximum and remained high through late time points in *ndt80* mutants (Fig. 3a, b and Extended Data Fig. 6). Heteroallele recombination was also increased (1.6-fold, Fig. 2b). Pachytene delay/arrest via Ndt80 inhibition is a hallmark of ZMM mutants<sup>9,24</sup>, suggesting that increased DSBs might be an indirect consequence of arrest<sup>22</sup>, perhaps analogous to increased DSB numbers when CHK-2 kinase activity is prolonged in *Caenorhabditis elegans*<sup>25,26</sup>. If so, then ZMM mutations should cause no change if Ndt80 is absent. However, more Spo11-oligo complexes (Fig. 3a, b and Extended Data Fig. 6) and heteroallele recombinants (Fig. 2b) accrued in *zip3 ndt80* and *msh5 ndt80* double mutants than in *ndt80* single mutants. Furthermore, *msh5*, *zip1* and *zip3* had similarly increased Spo11-oligo complexes (Fig. 1e) despite different arrest phenotypes (Fig. 3c). Thus, although the ZMM mutant DSB phenotype is probably influenced by the combined effects of Ndt80 inhibition and a hyperactivated DNA damage response, meiotic arrest per se does not explain ZMM mutant-provoked DSB increase.

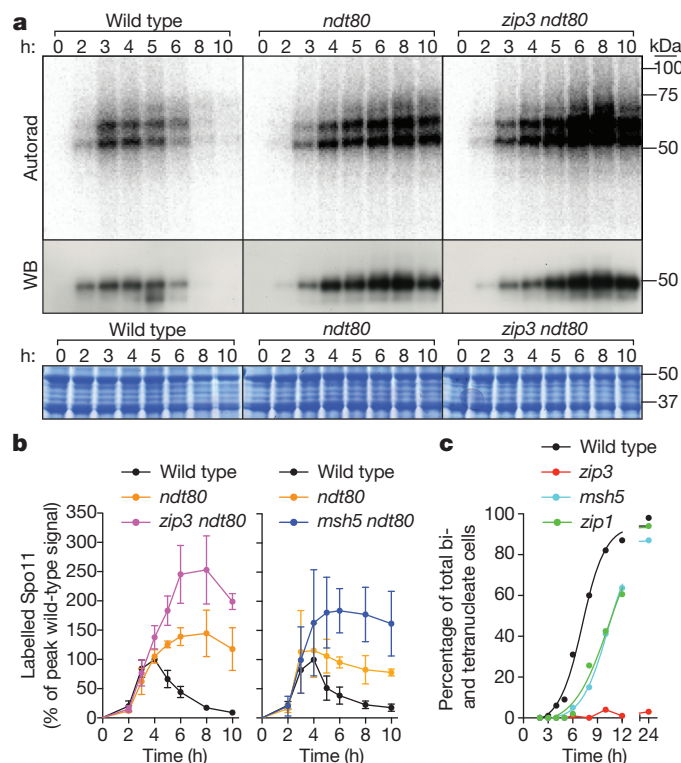
Instead, we infer that a ZMM-dependent process(es) is more directly responsible for inhibiting DSB formation. A plausible mechanism is that chromosomes that have engaged their homologues undergo structural changes that render them unfit Spo11 substrates<sup>8,27</sup>. 'Homologue engagement' could mean synaptonemal complex formation and/or progression of crossover-designated recombination intermediates, both promoted by ZMMs. Supporting this model, DSB-promoting factors Hop1 and Red1 accumulate on chromosomes in ZMM mutants<sup>28</sup>, proteins required for DSBs are displaced from pachytene chromosomes in wild-type yeast (for example, ref. 29), and Hop1 orthologues are displaced after synapsis in yeast and mouse<sup>27,28</sup>.

Hyper-rec behaviour in ZMM mutants is reconciled with tetrad data demonstrating globally reduced crossing over (for example, ref. 30) by noting that there is a reduced per-DSB likelihood of generating a crossover that offsets increased DSBs (Supplementary Discussion). Our findings





**Figure 2 | Hyper-rec phenotype of ZMM mutants.** **a**, Top, schematic of *arg4* heteroalleles, showing open reading frames and mutated restriction sites. Bottom, Spo11-oligo profile shows DSB distribution (RPM, reads per million mapped; smoothed with 201-base pair (bp) Hann window). **b**, Heteroallele recombination frequencies (mean  $\pm$  s.d.). \*Significantly different from wild type ( $P < 0.02$ , Student's *t*-test); \*\*significantly different from *ndt80* ( $P < 0.006$ ). **c**, Recombination reporter at the *CCT6* hotspot. **d**, Representative Southern blots of parental and recombinant DNA molecules (crossovers (COs) and noncrossover gene conversions (NCOs)) at *CCT6* resolved by two-dimensional gel electrophoresis. **e**, Recombination frequencies (mean  $\pm$  s.d.). Crossover frequencies were halved to convert to per-DSB equivalent because each crossover yields two recombinant molecules. \*Total recombination significantly different from wild type ( $P < 0.003$ ); \*\*crossing over significantly different from wild type ( $P < 0.04$ ). **f**, **g**, DSBs at *CCT6* and *GAT1*. A representative Southern blot probed for *CCT6* is in **f** and quantifications for *CCT6* and *GAT1* are in **g** (mean  $\pm$  s.d. for 3 cultures, except 8 h for *zip3* mutants at *GAT1*, analysed twice). JMs, joint molecules.



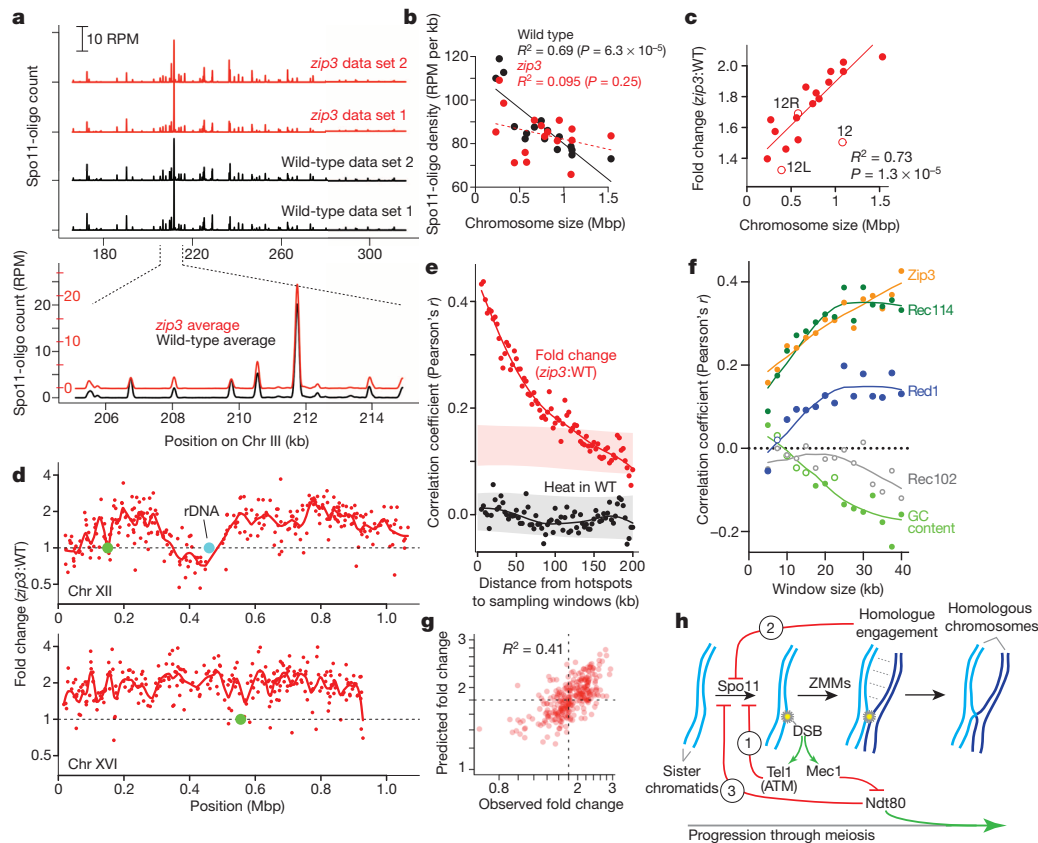
**Figure 3 | Separable effects of *ndt80* and ZMM mutations.** **a**, **b**, Spo11-oligo complex labelling from representative time courses is in **a** and quantification from  $\geq 3$  cultures (mean  $\pm$  s.d.) is in **b**. **c**, Meiotic progression (percentage of cells completing the first division).

can also be reconciled with studies that attributed persistent DSBs in ZMM mutants solely to increased DSB lifespan because, where tested, ZMM mutant and wild-type phenotypes were similar in *rad50* separation-of-function (*rad50S*) or *dmc1* backgrounds<sup>19,20,31–33</sup> (Extended Data Fig. 7a, b). Dmc1 is an essential strand exchange protein and *rad50S* mutants cannot remove Spo11 from DSB ends. As these mutations block recombination before ZMM proteins act, they are uninformative for querying ZMM mutant effects. This caveat may also apply to recombination-defective mutants in other organisms.

### Shaping the DSB landscape

If ZMM-dependent feedback works via chromosome structure changes linked to homologue engagement, then it should be spatially patterned. We tested this by deep-sequencing Spo11 oligos to map DSBs (Fig. 4a and Supplementary Table 3). Control cultures with a fully functional Spo11–protein A fusion agreed with each other and previous results<sup>11</sup> (Extended Data Figs 2, 8 and data not shown). The DSB ‘landscape’ is shaped by combinatorial action of many factors that operate hierarchically<sup>11,34</sup>. At short scales (sub-kilobase (kb)), the landscape is dominated by hotspots, mostly in nucleosome-depleted promoters. This pattern was unaffected in *zip3* mutants, in that DSBs formed in the same hotspots (Fig. 4a, Extended Data Figs 4f, 8c and Supplementary Table 4).

On larger scales, however, *zip3* showed substantial alterations. Smaller chromosomes form more crossovers per unit length than larger ones<sup>35</sup> because of variation in DSB levels<sup>11</sup>, but what controls DSB differences has been unclear. Remarkably, *zip3* mutation eliminated the normal inverse correlation between Spo11-oligo density and chromosome length (Fig. 4b). If the *zip3* map is scaled by 1.8-fold (on the basis of peak Spo11-oligo levels, Fig. 1e), all chromosomes had more DSBs but larger ones went up disproportionately (Fig. 4c). Thus, ZMM-dependent feedback is necessary for length-dependent recombination variation in wild type. Perhaps the number (not density) of DSBs governs speed or efficiency of homologue engagement: if so, smaller chromosomes might tend to have more time to accumulate DSBs. A nonexclusive possibility is that



**Figure 4 | Altered DSB distribution in *zip3* mutants.** **a**, Top, reproducibility of Spo11-oligo maps. Bottom, DSBs form at the same hotspots in *zip3* as wild type. Smoothed with 201-bp Hann window. **b**, Zip3 is required for chromosome size-dependent variation in Spo11-oligo density. Lines, least squares fits (dashed denotes nonsignificant). **c**, Larger chromosomes experience greater increase in Spo11 oligos. Fold change is the per-chromosome Spo11-oligo density in *zip3* over wild type (WT). Open circles, chromosome XII ('12', omitting rDNA length) and the portions of chromosome XII left or right of the rDNA ('12L', '12R'). Regression line treats 12L and 12R as separate chromosomes. **d**, Regional variation in response to *zip3* mutation. Each point is the change at a hotspot (plotted on log scale). Red lines, local regression (loess); green circles, centromeres. **e**, Local domains of correlated behaviour. Each point compares hotspots to their neighbours in 5-kb-wide windows the indicated distance away. Nearby hotspots show correlated behaviour for fold change in *zip3* (red), but not heat (Spo11-oligo

DSB suppression spreads far relative to chromosome length, with longer chromosomes providing more spreading room.

Subchromosomal domains differed in response to *zip3* mutation: Spo11-oligo frequencies increased less than average in 20-kb zones at telomeres and centromeres (where few DSBs form in wild type<sup>11,36,37</sup>), and were unchanged or reduced near the ribosomal DNA (rDNA), causing chromosome 12 to be an outlier in whole-chromosome analysis (Fig. 4c, d and Extended Data Fig. 9a). The remaining interstitial regions varied widely, with local regression along chromosomes suggesting alternating domains of greater or lesser change (Fig. 4d). Supporting this conclusion, the change in each hotspot correlated with the change in hotspots located nearby, with correlation strength decaying with distance (Fig. 4e).

To better understand these domains, we compared Spo11-oligo maps to chromosomal features including the distribution in wild type of Zip3 protein, chromosome structure proteins needed for normal DSBs (Hop1, Red1, Rec8) and proteins essential for Spo11 activity (Mei4, Mer2, Rec102, Rec104, Rec114), previously defined by chromatin immunoprecipitation (ChIP)<sup>38,39</sup>. The magnitude of change in Spo11-oligo density in *zip3* correlated with enrichment of Hop1, Rec114, Mei4, Mer2

frequency) in wild type (black). Shaded areas denote 95% confidence estimates for hotspots randomized within-chromosome (randomized  $r > 0$  for *zip3*-fold change because of the chromosome size effect). **f**, Correlation between log-fold change in *zip3* and binding of indicated proteins, binned in non-overlapping windows of varying size. For clarity, other proteins are in Extended Data Fig. 9b. Pericentric, subtelomeric and rDNA-proximal regions were censored. Closed symbols,  $P < 0.05$ . **g**, Fit of multiple regression model predicting changes (log scale) in Spo11-oligo density in 35-kb windows from ChIP data, G+C content and chromosome size (Supplementary Table 5). Dashed lines, observed mean fold change. **h**, Network of feedback circuits controlling DSB formation. Circuit 1: DSBs activate Tel1 (ATM in mouse), which inhibits further DSB formation. Circuit 2: ZMM-dependent interactions between homologous chromosomes inhibit Spo11. Circuit 3: Ndt80 shuts down DSB formation and drives pachytene exit; Mec1 kinase delays or blocks Ndt80 activation when DSBs are present.

and Red1, with highest correlation for binning windows  $\geq 20$  kb (Fig. 4f and Extended Data Fig. 9b, c, e). The distributions of these proteins are themselves correlated<sup>38</sup> (Extended Data Fig. 9d). We infer that large domains (tens of kb wide) enriched for these proteins tend to be more responsive to ZMM-dependent feedback. G+C content, Spo11-oligo density in wild type, and distributions of Rec8, Rec102 and Rec104 were uncorrelated or weakly anti-correlated when considered individually (Fig. 4f and Extended Data Fig. 9b, c, e, f). However, we observed a strong scale-dependent correlation with the distribution Zip3 displays when most DSBs have formed and homologues are engaging<sup>39</sup> (Fig. 4f and Extended Data Fig. 9c, e). Zip3 localizes to subsets of recombination sites<sup>17</sup>, so the positive correlation between Zip3 accumulation in wild type and altered DSB frequency in *zip3* suggests that Zip3 inhibits DSB formation, directly or indirectly, at sites of homologue engagement. Multiple regression indicates that these chromosomal features, plus chromosome size, explain  $\sim 40\%$  of the variation in *zip3*-induced DSB change (Fig. 4g and Supplementary Table 5). Our findings elucidate the locus-to-locus variability of ZMM mutant hyper-rec behaviour and reveal that ZMM-dependent feedback shapes the DSB landscape in wild type.

## Conclusions

We propose that the logic of DSB control involves a drive towards DSB formation that is restrained quantitatively, spatially and temporally by distinct but intersecting negative influences (Fig. 4h). We note several implications. First, Spo11 catalytic potential exceeds what is realized in any one meiosis. Thus, DSB numbers may underestimate severity of biochemical defects in mutants<sup>22</sup>. Second, counterintuitive effects arise when feedback loops are severed or hyperactivated, for example, in *dmc1* or *rad50S* backgrounds. The ZMM mutations likely impinge on multiple circuits simultaneously, removing restraints on Spo11 activity by disrupting homologue engagement and inhibiting Ndt80 activation, but also hyperactivating negative regulatory circuits via the DNA damage responsive kinase Tel1 (and possibly Mec1). This 'push-me-pull-you' interplay undoubtedly affects the final number and distribution of DSBs. Our results support the conclusion that crossovers in ZMM mutants are not identical in number and provenance to crossovers that form without ZMM intervention in wild type. Third, our findings explain puzzling aspects of *set1* mutant yeast and *Prdm9*<sup>-/-</sup> mutant mice. If DSB number control is separate from Spo11 targeting (which requires Set1 or PRDM9 (ref. 40)), then the default for Spo11 to make breaks until restrained by feedback explains why DSBs form in relatively normal numbers but different locations in these mutants. This also undermines more extreme versions of the 'hotspot paradox' in which biased gene conversion is predicted to eliminate all hotspots over time and thereby prevent DSB formation (for example, ref. 41): the logic of DSB control makes it impossible for inactivation of individual hotspots to render chromosomes immune to Spo11. Fourth, our findings support the hypothesis that altered DSB distributions tied to feedback control are the source of altered recombination distributions caused by certain mutations or heterozygosity for large-scale chromosome structure variants in other species<sup>6,7</sup>. Finally, we speculate that organisms such as mouse readily form synaptonemal complexes between nonhomologous chromosome segments late in prophase (for example, ref. 8) as a means to shut down unproductive DSB formation in karyotypically unbalanced cells.

Our findings provide a holistic view of DSB control in the broader context of meiotic chromosome dynamics and meiotic progression, and explain how DSB formation is homeostatic and therefore robust against cell-to-cell variation, environmental perturbation, and chromosome variants encountered in outcrosses.

## METHODS SUMMARY

Yeast strains are of the SK1 background (Supplementary Table 6). Synchronized meiotic cultures were prepared according to standard methods. Labelling of Spo11-oligo complexes and purification, amplification and sequencing of Spo11 oligos were carried out using methods adapted from previous studies<sup>10,11</sup>.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 14 July 2013; accepted 31 January 2014.**

**Published online 6 April 2014.**

- Keeney, S. Mechanism and control of meiotic recombination initiation. *Curr. Top. Dev. Biol.* **52**, 1–53 (2001).
- Lange, J. et al. ATM controls meiotic double-strand-break formation. *Nature* **479**, 237–240 (2011).
- Joyce, E. F. et al. *Drosophila* ATM and ATR have distinct activities in the regulation of meiotic DNA damage and repair. *J. Cell Biol.* **195**, 359–367 (2011).
- Zhang, L., Kim, K. P., Kleckner, N. E. & Storlazzi, A. Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1/ATR and Tel1/ATM, once per quartet of chromatids. *Proc. Natl Acad. Sci. USA* **108**, 20036–20041 (2011).
- Carballo, J. A. et al. Budding yeast ATM/ATR control meiotic double-strand break (DSB) levels by down-regulating Rec114, an essential component of the DSB-machinery. *PLoS Genet.* **9**, e1003545 (2013).
- Bhagat, R., Manheim, E. A., Sherizen, D. E. & McKim, K. S. Studies on crossover-specific mutants and the distribution of crossing over in *Drosophila* females. *Cytogenet. Genome Res.* **107**, 160–171 (2004).
- Henzel, J. V. et al. An asymmetric chromosome pair undergoes synaptic adjustment and crossover redistribution during *Caenorhabditis elegans* meiosis: implications for sex chromosome evolution. *Genetics* **187**, 685–699 (2011).
- Kauppi, L. et al. Numerical constraints and feedback control of double-strand breaks in mouse meiosis. *Genes Dev.* **27**, 873–886 (2013).
- Lynn, A., Soucek, R. & Börner, G. V. ZMM proteins during meiosis: crossover artists at work. *Chromosome Res.* **15**, 591–605 (2007).
- Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053–1057 (2005).
- Pan, J. et al. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **144**, 719–731 (2011).
- Xu, L., Ajimura, M., Padmore, R., Klein, C. & Kleckner, N. NDT80, a meiosis-specific gene required for exit from pachytene in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**, 6572–6581 (1995).
- Sym, M., Engebrecht, J. A. & Roeder, G. S. ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell* **72**, 365–378 (1993).
- Ross-Macdonald, P. & Roeder, G. S. Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* **79**, 1069–1080 (1994).
- Hollingsworth, N. M., Ponte, L. & Halsey, C. MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev.* **9**, 1728–1739 (1995).
- Chua, P. R. & Roeder, G. S. Zip2, a meiosis-specific protein required for the initiation of chromosome synapsis. *Cell* **93**, 349–359 (1998).
- Agarwal, S. & Roeder, G. S. Zip3 provides a link between recombination enzymes and synaptonemal complex proteins. *Cell* **102**, 245–255 (2000).
- Tsubouchi, T., Zhao, H. & Roeder, G. S. The meiosis-specific zip4 protein regulates crossover distribution by promoting synaptonemal complex formation together with zip2. *Dev. Cell* **10**, 809–819 (2006).
- Shinohara, M., Oh, S. D., Hunter, N. & Shinohara, A. Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. *Nature Genet.* **40**, 299–309 (2008).
- Börner, G. V., Kleckner, N. & Hunter, N. Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* **117**, 29–45 (2004).
- Allers, T. & Lichten, M. Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106**, 47–57 (2001).
- Gray, S., Allison, R. M., Garcia, V., Goldman, A. S. & Neale, M. J. Positive regulation of meiotic DNA double-strand break formation by activation of the DNA damage checkpoint kinase Mec1(ATR). *Open Biol.* **3**, 130019 (2013).
- Rockmill, B. et al. High throughput sequencing reveals alterations in the recombination signatures with diminishing Spo11 activity. *PLoS Genet.* **9**, e1003932 (2013).
- Tung, K. S., Hong, E. J. & Roeder, G. S. The pachytene checkpoint prevents accumulation and phosphorylation of the meiosis-specific transcription factor Ndt80. *Proc. Natl Acad. Sci. USA* **97**, 12187–12192 (2000).
- Rosu, S. et al. The *C. elegans* DSB-2 protein reveals a regulatory network that controls competence for meiotic DSB formation and promotes crossover assurance. *PLoS Genet.* **9**, e1003674 (2013).
- Stamper, E. L. et al. Identification of DSB-1, a protein required for initiation of meiotic recombination in *Caenorhabditis elegans*, illuminates a crossover assurance checkpoint. *PLoS Genet.* **9**, e1003679 (2013).
- Wojtasz, L. et al. Mouse HORMAD1 and HORMAD2, two conserved meiotic chromosomal proteins, are depleted from synapsed chromosome axes with the help of TRIP13 AAA-ATPase. *PLoS Genet.* **5**, e1000702 (2009).
- Smith, A. V. & Roeder, G. S. The yeast Red1 protein localizes to the cores of meiotic chromosomes. *J. Cell Biol.* **136**, 957–967 (1997).
- Kee, K., Protacio, R. U., Arora, C. & Keeney, S. Spatial organization and dynamics of the association of Rec102 and Rec104 with meiotic chromosomes. *EMBO J.* **23**, 1815–1824 (2004).
- Chen, S. Y. et al. Global analysis of the meiotic crossover landscape. *Dev. Cell* **15**, 401–415 (2008).
- Nakagawa, T. & Ogawa, H. The *Saccharomyces cerevisiae* MER3 gene, encoding a novel helicase-like protein, is required for crossover control in meiosis. *EMBO J.* **18**, 5714–5723 (1999).
- Jessop, L., Rockmill, B., Roeder, G. S. & Lichten, M. Meiotic chromosome synapsis-promoting proteins antagonize the anti-crossover activity of sgs1. *PLoS Genet.* **2**, e155 (2006).
- Falk, J. E., Chan, A. C., Hoffmann, E. & Hochwagen, A. A Mec1- and PP4-dependent checkpoint couples centromere pairing to meiotic recombination. *Dev. Cell* **19**, 599–611 (2010).
- Lichten, M. in *Recombination and Meiosis: Models, Means, and Evolution* Vol. 3 (eds R. Egel & D. H. Lankenau) 165–193 (Springer-Verlag, 2008).
- Kaback, D. B., Guacci, V., Barber, D. & Mahon, J. W. Chromosome size-dependent control of meiotic recombination. *Science* **256**, 228–232 (1992).
- Blitzblau, H. G., Bell, G. W., Rodriguez, J., Bell, S. P. & Hochwagen, A. Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres. *Curr. Biol.* **17**, 2003–2012 (2007).
- Buhler, C., Borde, V. & Lichten, M. Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. *PLoS Biol.* **5**, e324 (2007).
- Panizza, S. et al. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* **146**, 372–383 (2011).
- Serrentino, M. E., Chaplais, E., Sommermeyer, V. & Borde, V. Differential association of the conserved SUMO ligase Zip3 with meiotic double-strand break sites reveals regional variations in the outcome of meiotic recombination. *PLoS Genet.* **9**, e1003416 (2013).



40. Borde, V. & de Massy, B. Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. *Curr. Opin. Genet. Dev.* **23**, 147–155 (2013).
41. Pineda-Krch, M. & Redfield, R. J. Persistence and loss of meiotic recombination hotspots. *Genetics* **169**, 2319–2333 (2005).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We are grateful to S. Burgess, N. Hunter, N. Kleckner, K. Ohta, M. Rout and A. Shinohara for strains or epitope tagging constructs; F. Klein for sharing data; S. Shuman for gifts of T4 RNA ligase; A. Viale and the Memorial Sloan Kettering Cancer Center (MSKCC) Genomics Core Laboratory for sequencing; and N. Socci and the MSKCC Bioinformatics Core for assistance mapping Spo11 oligos. This work was

supported by National Institutes of Health grant R01 GM058673. S.K. is an Investigator of the Howard Hughes Medical Institute.

**Author Contributions** D.T., N.M. and X.Z. performed experiments and analysed the data. N.M. optimized purification of the Spo11–protein A fusion, prepared the sequencing libraries to map DSBs and performed preliminary data analysis on sequencing reads. S.K. analysed sequencing data. D.T. and S.K. wrote the paper.

**Author Information** Sequencing data were deposited at GEO under accession number GSE48299. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.K. ([s-keeney@ski.mskcc.org](mailto:s-keeney@ski.mskcc.org)).

## METHODS

**Yeast strains and plasmids.** Strains were of the SK1 background (Supplementary Table 6). The *zip1* deletion and one of the *ndt80* deletions (*ndt80Δ::LEU2*) were provided by N. Kleckner, the *msh5* and *zip3* deletions were provided by N. Hunter, the *spo16* deletion was provided by A. Shinohara, and a second *ndt80* deletion (*ndt80Δ::kanMX4*) was provided by S. Burgess. The *dmc1*, *pph3* and *zip4* deletions were made by replacing the coding sequence with the hygromycin B phosphotransferase gene (*hphMX4*). Gene disruption was verified by PCR. All mutants analysed were moved into the desired tester strain backgrounds by crossing and tetrad dissection. The *SPO11-Flag* strain was provided by K. Ohta and the protein A tagging construct was provided by M. Rout. The constructs for two-dimensional gel electrophoresis analysis of crossover and noncrossover recombinants at *CCT6*, *ERG1* and *GAT1* were engineered by a series of two-step gene replacements. For *CCT6* on chromosome IV, *Sall* sites were introduced in intergenic regions at Saccharomyces Genome Database (SGD) coordinates 832534 and 838251 in one strain; *Sall* sites were introduced separately at coordinates 833537 (in *YDR186c*) and 837893 (in *CCT6*) in another strain along with a *SmaI* site between *YDR186c* and *CCT6* at coordinates 835802 and 835803. For *ERG1* on chromosome VII, *SacII* sites were introduced in intragenic regions at coordinates 844276 (in *RBG2*) and 854464 (in *OKPI*) along with a *Sall* site at coordinate 848724 (between *ERG1* and *ATP2*). In a separate strain, *SacII* sites were introduced at coordinate 845470 (intergenic) and coordinate 852145 (in *PBP1*). For *GAT1* on chromosome VI, *KpnI* sites were introduced at coordinates 90967 and 100083 (both intergenic) along with a *BamHI* site between *FRS2* and *GAT1* at coordinates 95715 and 95717. Separately, *KpnI* sites were introduced at coordinates 92986 (*BUD27*) and coordinate 98899 (intergenic). Further details are in ref. 42 and available on request.

**Culture methods.** With the exception of Spo11-oligo mapping, synchronous meiotic cultures were prepared as described<sup>43,44</sup>. In brief, cells were grown in YPA (1% yeast extract, 2% Bacto Peptone, 1% potassium acetate) for 13.5–14 h at 30 °C, collected, re-suspended in 2% potassium acetate, and sporulated at 30 °C.

**Meiotic division profiles.** Aliquots were collected from synchronous meiotic cultures and fixed in 47.5% (v/v) ethanol and 0.05 μg ml<sup>-1</sup> DAPI. Mono-, bi- and tetranucleate cells were scored by epifluorescence microscopy.

**Direct DSB measurements and heteroallele recombination analysis.** High-molecular-weight DNA was prepared and separated by pulsed-field gel electrophoresis as described<sup>45</sup>. DNA was probed with part of the *CHA1* open reading frame (SGD coordinates 15838 to 16857), *SKI8* (coordinates 90062 to 91228), *YHL042w* (coordinates 15671 to 16112) or *POT1* (coordinates 40223 to 40728). DSB analyses at *CCT6*, *ERG1* and *GAT1* were performed as described<sup>41</sup>. Blots were quantified by phosphorimager. For the *DMC1*<sup>+</sup> pulsed-field DSB analyses, the signal above the parental band (including the well) was split between the parental and DSB signals. For quantification in Fig. 1c and Extended Data Fig. 7b, our main interest was absolute DSB levels, so we calculated the average number of DSBs per chromatid assuming a Poisson distribution of breaks among and along chromatids in the population:  $P(n) = (\mu^n e^{-\mu})/n!$  (where  $\mu$  is the mean number of DSBs per chromatid and  $P(n)$  is the probability that  $n$  DSBs occur on a single chromatid). The observed parental-length signal ( $U_{\text{obs}}$ , for 'unbroken') approximates the true unbroken fraction (that is,  $P(0) \approx U_{\text{obs}}$ ), so the mean total number of DSBs per chromatid in the population ( $DSB_{\text{total}}$ ) can be estimated as  $-\ln(U_{\text{obs}})$ . This calculation helps correct for multiple DSBs on the same chromatid. A full description of the method, and confirmation that it does not overestimate DSB numbers, will be provided elsewhere (H. Murakami & S.K., unpublished observations).

Because relative DSB levels were our main interest for data in Extended Data Fig. 2, we did not apply a Poisson correction and expressed DSBs instead as detectable broken DNA as per cent of total DNA in the lane. For the locus-specific DSB analyses the signal between the parental band and the wells was measured and apportioned evenly between the parental and DSB values. The frequency of meiotic recombination at *ARG4* was determined by heteroallele recombination analysis as described<sup>46</sup>.

**Two-dimensional gel electrophoresis analysis of crossovers and noncrossovers.** Cultures were grown and sporulated as described above at 30 °C. Samples (15 ml) were collected at 10 h and washed twice with 5 ml of 50 mM EDTA, pH 8.0. For analyses at *CCT6* and *ERG1*, DNA was prepared in agarose plugs as described<sup>45</sup>. For analysis at *GAT1*, DNA was prepared for conventional agarose electrophoresis<sup>47</sup>. DNA embedded in agarose plugs was digested with the appropriate restriction endonuclease (*CCT6*, *Sall*; *ERG1*, *SacII*), then electrophoresed at room temperature (20–23 °C) for 24–26 h at 1.7 V cm<sup>-1</sup> on 0.5% agarose in 0.5 × TBE (Tris/borate/EDTA). DNA prepared for conventional agarose electrophoresis (*GAT1*) was digested with *KpnI*, then electrophoresed at room temperature for 24 h at 1.7 V cm<sup>-1</sup> on 0.5% agarose in 1 × TBE. A ~10.5 cm gel slice containing the region of interest was then excised. For analysis at *CCT6* and *ERG1* the gel slice was washed twice in the appropriate NEBuffer supplemented with 20 μg ml<sup>-1</sup> BSA. For analysis at *GAT1* the gel slice was washed twice in 10 mM Tris-HCl, pH 8.1, followed by one

wash in NEBuffer 4 supplemented with 100 μg ml<sup>-1</sup> BSA. Liquid was then replaced with fresh NEBuffer supplemented with 20 μg ml<sup>-1</sup> BSA or 100 μg ml<sup>-1</sup> BSA, then 4,900–5,000 units of the appropriate restriction endonuclease (*CCT6*, *SmaI*; *GAT1*, *BamHI*; *ERG1*, *Sall*) was added and incubated first at 4 °C overnight then at the optimal incubation temperature for ~24 h. The gel slice was then cast in a 0.6% agarose gel in 0.5 × TBE (*CCT6* and *ERG1*) or 1 × TBE (*GAT1*), then electrophoresed perpendicular to the first dimension at ~1.3 V cm<sup>-1</sup> for ~20 h at room temperature. DNA was probed with part of the *CCT6* open reading frame (SGD coordinates 837413 to 837865), part of the *GAT1* open reading frame (SGD coordinates 95968 to 97490 or SGD coordinates 96500 to 97491), or part of the *PBP1* open reading frame (SGD coordinates 851379 to 851869).

**End-labelling of Spo11-oligo complexes and western blot analysis.** Lysates and extracts were prepared as previously described<sup>48</sup>. Immunoprecipitation of Spo11-oligo complexes was performed using 5 μg of mouse monoclonal anti-Flag M2 antibody (Sigma). Precipitated Spo11-oligo complexes were end-labelled in NEBuffer 4 (New England Biolabs) containing 3–10 μCi of [ $\alpha$ -<sup>32</sup>P]dCTP and terminal deoxynucleotidyl transferase (TdT)<sup>48</sup>. 25 μl of reaction mixture was added to the beads, mixed, and incubated at 37 °C for 1–2 h. Spo11-oligo complexes were eluted by adding 25 μl of NUPAGE loading buffer (diluted to 2× and supplemented with 83.3 mM dithiothreitol (Invitrogen)) and boiling for 5 min. End-labelled Spo11-oligo complexes were separated on a Novex 4–12% gradient denaturing polyacrylamide gel (Invitrogen) then transferred onto PVDF membrane using the iBlot protocol (Invitrogen) and visualized by phosphorimager. Blots were probed with mouse monoclonal anti-Flag M2 conjugated to horseradish peroxidase (Sigma). Chemiluminescent detection was performed according to the manufacturer's instructions (ECL+ or ECL Prime, Amersham). Protein quantity was estimated by separating 1 μl of extract on a Novex 4–12% gradient denaturing polyacrylamide gel and staining with Coomassie blue.

**Spo11-oligo purification for mapping.** Spo11 oligos were prepared for sequencing similar to methods described previously<sup>11</sup>, with modifications. Haploid strains with Spo11 C-terminally tagged with five copies of the protein A tag were patched from a frozen stock onto a YP-glycerol plate and grown at 30 °C overnight to select for respiration competence. Cells were mated on YPD (1% yeast extract, 2% Bacto Peptone, 2% dextrose) plates then streaked for single colonies and grown for 48 h at 30 °C. A single colony was inoculated into 5 ml liquid YPD medium and grown overnight at 30 °C. The saturated YPD culture was used to inoculate 25 ml liquid SPS medium (0.5% yeast extract, 1% peptone, 0.67% yeast nitrogen base without amino acids (Difco), 1% potassium acetate, 0.05 M potassium biphthalate, pH 5.5) to *OD*<sub>600</sub> 0.8 and grown for 7 h at 30 °C. This culture was used to inoculate 1 l SPS medium in a 2.8 l baffled Fernbach flask to *OD*<sub>600</sub> 0.05. Flasks were incubated at 30 °C for 12–16 h, to *OD*<sub>600</sub> 4.5–6. Cells were collected by centrifugation, washed once in deionized water, re-suspended in 0.6 l sporulation medium (2% potassium acetate and 0.001% antifoam 204) and incubated in 2.8 l baffled flasks (0.6 l per flask) at 30 °C for 4 h (wild type) or 5 h (*zip3*) to approximate times of peak Spo11-oligo levels (Fig. 1e).

Cells were centrifuged and washed with 50 mM EDTA, transferred to a 30-ml syringe, extruded into liquid nitrogen, and stored at –80 °C. Yeast cell powder was prepared by placing the frozen paste into canisters of a Retsch MM301 mill (pre-chilled in liquid nitrogen) and grinding five times for 3 min at 30 Hz. Yeast powder was transferred to a pre-chilled 50-ml tube and stored at –80 °C. Extract was prepared by transferring the yeast powder to a pre-chilled 40-ml glass Dounce homogenizer and homogenizing in two volumes of cold 10% trichloroacetic acid. Lysate was centrifuged at 14,000 r.p.m. in an SS-34 rotor (Sorvall) for 20 min. The supernatant was removed and cell pellet was re-suspended in SDS extraction buffer (2% SDS, 0.5 M Tris-HCl, pH 8.1, 10 mM EDTA, 0.005% bromophenol blue). β-mercaptoethanol was added to 0.288 M, the extract boiled in a water bath for 5 min, then centrifuged at 14,000 r.p.m. in an SS-34 rotor for 20 min.

The supernatant was poured into fresh tubes, diluted with an equal volume of 2× immunoprecipitation (IP) buffer (2% Triton X-100, 30 mM Tris-HCl, pH 8.0, 300 mM NaCl, 2 mM EDTA) and incubated with CL6B-sepharose beads (GE) for mock IP (4 h at 4 °C mixing end-over-end, 1.5 ml extract per 200 μl beads). Supernatant was removed into fresh tubes and mock beads were stored on ice. The supernatant was incubated with 200 μl IgG Sepharose Fast Flow beads (GE) per 1.5 ml of extract for 4 h at 4 °C mixing end-over-end, then beads were recovered. Mock and IP beads were washed 3 times with 10 ml cold 1 × IP buffer. Protein was eluted from mock or IP beads with 350 μl 2 × NuPAGE LDS buffer (Invitrogen) by boiling for 5 min, followed by a second elution with 350 μl 0.5 × NuPAGE LDS buffer. The eluates were combined and diluted with 700 μl of 2 × IP buffer, then incubated with 200 μl fresh CL6B-sepharose beads (mock) or IgG Sepharose Fast Flow beads (IP), 4 °C overnight with end-over-end rotation. The beads were recovered and subsequently washed with 1 ml Proteinase K buffer (100 mM Tris-HCl, pH 7.4, 1 mM EDTA, 0.5% SDS, 1 mM CaCl<sub>2</sub>) lacking SDS, then re-suspended in 600 μl Proteinase K buffer and 100 μg purified Proteinase K, and incubated overnight

at 50 °C with end-over-end rotation. The supernatant was collected using a SPIN-X tube (Corning) and ethanol precipitated with 0.3 volume of 9 M ammonium acetate, 10 µg of DNA-free glycogen and 2.5 volumes of 100% ethanol. Spo11 oligos were quantified by end labelling with [ $\alpha$ -<sup>32</sup>P]GTP and TdT (Fermentas) and comparing to a known quantity of similarly labelled 30-nucleotide synthetic oligo<sup>11</sup>.

**Library preparation for sequencing.** Approximately 300 fmol of Spo11 oligos were subjected to GTP tailing at their 3' ends. Material eluted from mock beads was processed in parallel to determine specificity of the IP (data not shown). Tailing was carried out in a total volume of 40 µl containing 1 × NEBuffer 4, 20 U TdT and 13.8 µM GTP at 37 °C for 5 h, followed by heat inactivation of TdT at 75 °C for 10 min. The tailed oligos were ligated to a double-stranded DNA adaptor optimized for the Illumina HiSeq platform as follows: the tailing reaction was supplemented by addition of 10 × T4 RNA ligase 2 buffer (500 mM Tris-HCl, pH 7.6, 50 mM MgCl<sub>2</sub>, 50 mM β-mercaptoethanol) to 1 ×, 25 mM ATP to 0.5 mM, 5 pmol double-stranded customized P7 adaptor, 300 fmol T4 RNA ligase 2 (gift from Stewart Shuman, MSKCC), and dH<sub>2</sub>O to a final volume of 50 µl. P7 adaptor sequences are 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCCC and 5'-PAGATCGGAAGAGCACACGTCTGAACTCCAGTCACppT, where CppT is an inverted 3'-3' linkage to block ligation. These oligos were annealed and purified by non-denaturing polyacrylamide gel electrophoresis before use in the ligation reaction. Ligation was carried out overnight at room temperature. Complementary strands of Spo11 oligos were synthesized as follows: the ligation reaction was supplemented with 2 mM dNTP to a final concentration of 30 µM and 10 U Klenow polymerase (New England Biolabs), and incubated at 25 °C for 15 min. After Klenow inactivation (75 °C, 10 min), extension reactions were supplemented with 0.3 volume of 9 M ammonium acetate, 10 µg of DNA-free glycogen and 2.5 volumes of 100% ethanol. DNA was precipitated at -20 °C overnight and centrifuged at 16,000g. The pellet was rinsed with 70% ethanol, air dried, and dissolved in a mixture of 9 µl water and 15 µl formamide loading buffer. Extension products and 10-bp ladder (radiolabelled with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP) were separated on a 10% denaturing polyacrylamide gel. The region between ~55–200 nt (equivalent to ~10–50-nucleotide Spo11 oligos with (rG)<sub>3–5</sub> tails plus ligated adaptor) was excised, crushed, and eluted in 400 µl 10 mM Tris-HCl, pH 8.0 at 37 °C overnight with mixing. Elution mixture was spun through a SPIN-X tube, then 0.3 volume of 9 M ammonium acetate, 10 µg DNA-free glycogen, and 2.5 volume of 100% ethanol were added. DNA was precipitated on dry ice at -20 °C overnight and centrifuged at 16,000g. Pellet was rinsed with 70% ethanol and air dried. The 3' ends of gel-purified, denatured DNA strands were tailed with GTP by dissolving the dried pellet in 40 µl tailing reaction containing 1 × NEBuffer 4, 30 U TdT and 50 µM GTP, then incubating at 37 °C for 5 h. The tailed oligos were ligated to a second set of customized double-stranded DNA adaptors (P5) and complementary strands were synthesized as above. The P5 adaptor is a mixture of four duplexes. The oligos for one duplex are 5'-ACACTCTTCCCTACACGACGCTCTTCCGATCTAGTCTCCCC (top strand) and 5'-PAGACTAGATCGGAAGAGCGTCGTGTAGG GAAAGAGTGTppT. The four different P5 duplexes have different sequences at the underlined positions (top strand: AGTC, GTCA, TCAG, CAGT, respectively), which are the first four bases that will be read by the sequencer. Complementary pairs of oligos were annealed separately and purified by non-denaturing electrophoresis, then the four duplexes were mixed in approximately equimolar ratio. This provides diversity of base composition at the beginning of the sequencing reaction. If this diversifier region were not present, the sequencer would encounter a homogeneous oligo-C sequence for every read, compromising ability to detect individual amplification clusters immobilized on the flow cell surface.

To estimate the yield, test PCR was carried out in a total of 30 µl containing 1–2% of the final Klenow extension reaction, 1 × PCR buffer (Invitrogen), 2 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 1.5 U Taq polymerase (Invitrogen) and 1 µM P5 primer: 5'-AATGATACGGGACACCGAGATCTACACTCTTCCCTACACGACG, comprising Illumina sequencing primer and part of P5 adaptor, 1 µM Indexing primer: 5'-CAAGCAGAAGACGGCATACGAGATCGTGTGACTGGAGTTCAGACGTGTG (underlined sequence is Illumina HiSeq Index 1, replaced as appropriate with the sequence of other indices according to manufacturer instructions), comprising Illumina-specific primer and part of the P7 adaptor). The mixture was divided into three tubes and PCR was initiated by a denaturation step at 94 °C for 10 s, followed by 20 cycles of amplification (94 °C for 10 s, 60 °C for 10 s and 72 °C for 10 s). PCRs were combined and the products were electrophoresed on a 10% non-denaturing polyacrylamide gel with low-molecular-weight DNA ladder (New England Biolabs) to determine the size and quantity of PCR product after staining with ethidium bromide. Template for Illumina sequencing was prepared by a large-scale PCR with same conditions as above (but only 16 cycles instead of 20), scaled up to a total volume of 640 µl containing the desired amount of the Klenow-extended products. Sequencing was performed using Illumina HiSeq in the Memorial Sloan Kettering Cancer Center Genomics Core Laboratory.

**Bioinformatic analysis.** Statistical analyses were performed using R version 2.15.3 (<http://www.r-project.org/>)<sup>49</sup> or GraphPad Prism 6.0. Mapping of Illumina reads to the target genome was performed using a pipeline essentially as described<sup>11</sup>. In brief, adaptor sequences were removed from both the 5' and 3' ends, then reads were mapped to the *S. cerevisiae* genome (SGD version June 2008, that is, sacCer2) using *gmap*-ls (2\_1\_1b) from the SHRiMP mapping package<sup>50</sup>. The specific mapping parameters used were -U -g -1000 -q -1000 -m 10 -i -20 -h 100 -r 50% -o 1001, which forces ungapped alignments (-U by itself did not suppress all gapped alignments so we set an effectively infinite gap opening penalty). To increase sensitivity for short reads we set the seeds to the following: -n 1 -s 1111111111, 11110111101111, 1111011100100001111, 1111000011001101111. After mapping, the reads were separated into unique and multiple mapping sets, but only uniquely mapping reads were analysed in this study (multiple mapping reads constituted a small minority of the total). A full copy of the source code is available online at [http://cbio.mskcc.org/public/Thacker\\_ZMM\\_feedback](http://cbio.mskcc.org/public/Thacker_ZMM_feedback).

Because the rDNA array is represented in the SGD assembly by only 1.9 copies of the repeat unit, oligos that span the boundary between repeats map to a single position even though they come from a repetitive sequence. Therefore, reads of this type were moved to the multiple mapping set. The wild-type data sets contained a small number of spurious reads (<0.4% of total) from contamination of the Spo11-oligo sequencing libraries with PCR primers from the *TEL1* locus; these reads were deleted from the maps. Because of the variable number of rG residues added by terminal transferase to the 3' end of Spo11 oligos and to the 3' end of the reverse complementary strands, there is ambiguity in defining the precise start and end positions for reads that map to positions starting with one or more C residues or ending in one or more G residues<sup>11</sup>. In such cases, the 5' and 3' ends of each read were defined so as to provide the longest contiguous sequence match with the genome.

Raw and processed sequence reads have been deposited in the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) (accession number GSE48299). This accession also contains the curated maps (unique mapping reads only) in wiggle format to allow direct visualization in appropriate genome browsers, for example, the UCSC browser (<http://genome.ucsc.edu/>) using genome version sacCer2).

For the studies here, our focus is on the number and position of DSBs rather than characteristics of the Spo11 oligos themselves, so maps were distilled to record just the positions of 5' ends of oligos. Each map was normalized to the total number of reads that mapped uniquely to a chromosome (RPM; excluding reads mapping to mitochondrial DNA or the 2µ plasmid), then wild-type and *zip3* maps were averaged. Normalized Spo11-oligo counts within the 3,600 previously identified hotspots<sup>11</sup> are compiled for each data set in Supplementary Table 4. In analyses evaluating the fold change (that is, Fig. 4c, d, g and Extended Data Fig. 9a, d–f), we assumed a global increase in Spo11-oligo number of 1.8-fold based on the difference in peak steady-state levels (Fig. 1e). To prevent dividing by zero and to minimize variability of ratios caused by small changes in denominators, we added a small constant to numerator and denominator before taking the ratio (20 RPM for hotspot-based ratios (approximately 15% of median hotspot Spo11-oligo count), or 0.1 RPM per kb for bin-based ratios (0.13% of median Spo11-oligo density per bin)). Where indicated, Spo11-oligo maps were smoothed with a 201-bp Hann window.

For the correlation analysis in Fig. 4f and Extended Data Fig. 9b, chromosomes were divided into non-overlapping bins of the indicated sizes. Bins that overlapped censored regions (within 20 kb of telomeres, within ± 10 kb of centromeres, or in the region from 60 kb leftward to 30 kb rightward of the rDNA) were discarded. The published ChIP enrichment data (log<sub>2</sub> of ChIP/input; from GEO accession GSE29860 (ref. 38) or Supplementary Table 3 from ref. 39) were averaged within each bin, then compared to the mean log-fold change in Spo11-oligo density in *zip3* and correlation coefficients were calculated. The log-transformed data were approximately normally distributed so we used Pearson's *r*, but similar overall patterns were obtained if we used Kendall's tau (data not shown).

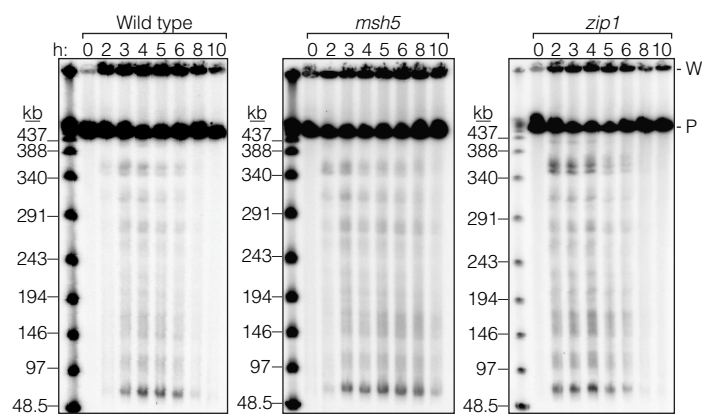
Multiple linear regression was performed using the 'lm' function in R. Data were averaged in non-overlapping bins of 35 kb, censored for subtelomeric, pericentric and rDNA-proximal regions as described above. Principal component analysis was performed on the correlation matrix of the Rec114, Mei4, Mer2, Hop1 and Red1 ChIP data using the 'princomp' function in R. The first principal component accounted for 92.7% of the variance in this data set; the remaining principal components were discarded as they accounted for only 4.1%, 2.2%, 0.8% and 0.3% of the variance, respectively.

To assess spatial correlations for the change in Spo11-oligo density (Fig. 4e), we calculated the correlation coefficient (Pearson's *r*) between the log-fold change at DSB hotspots and the log-fold change for hotspots located within a set of 5-kb windows centred a distance *D* to the right of each hotspot centre. We varied *D* from 5 to 200 kb in steps of 2.5 kb and calculated the correlation coefficient separately for

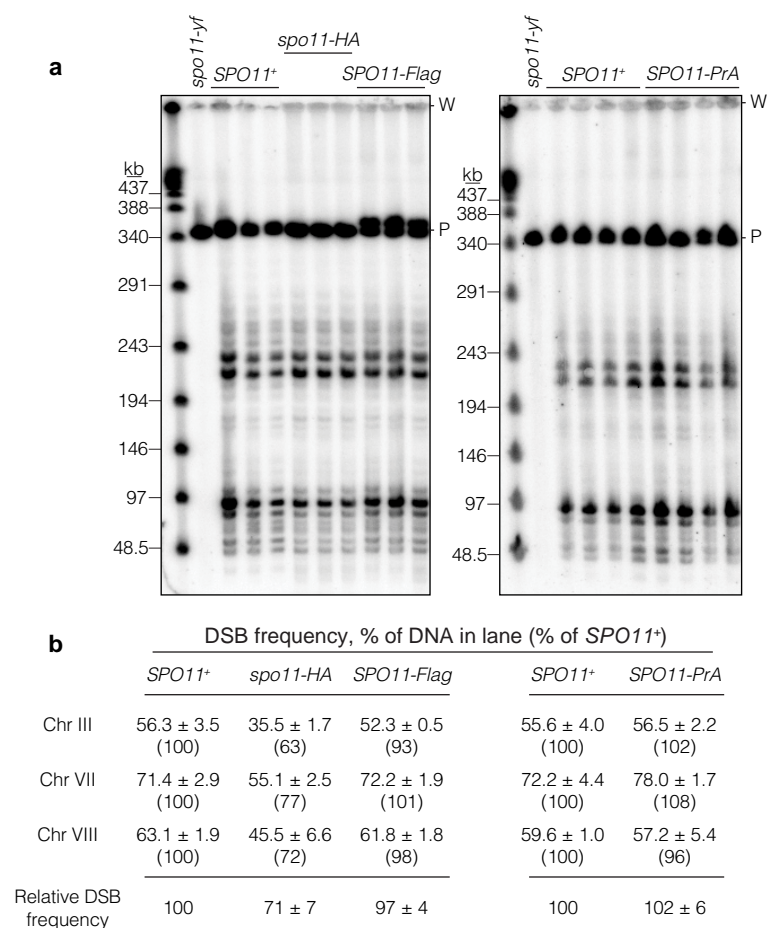


each distance. For comparison, we performed the same analysis to evaluate the correlation between absolute heat (log of the Spo11-oligo count) in hotspots and the heat in 5-kb windows at varying distances. To generate randomized controls for this analysis, we randomly reassigned the heats or log-fold change between hotspots within a chromosome. This randomization strategy preserves the non-random placement of hotspots relative to one another and preserves the correlated behaviour (if any) across whole chromosomes. Randomization was repeated 100 times to provide the estimates of the 95% confidence intervals shown in the Fig. 4e.

42. Thacker, D. *Meiotic Recombination in Saccharomyces Cerevisiae: from New Assays to New Insights*. PhD thesis, Cornell Univ. (2012).
43. Alani, E., Padmore, R. & Kleckner, N. Analysis of wild-type and *rad50* mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. *Cell* **61**, 419–436 (1990).
44. Padmore, R., Cao, L. & Kleckner, N. Temporal comparison of recombination and synaptonemal complex formation during meiosis in *S. cerevisiae*. *Cell* **66**, 1239–1256 (1991).
45. Borde, V., Goldman, A. S. & Lichten, M. Direct coupling between meiotic DNA replication and recombination initiation. *Science* **290**, 806–809 (2000).
46. Diaz, R. L., Alcid, A. D., Berger, J. M. & Keeney, S. Identification of residues in yeast Spo11p critical for meiotic DNA double-strand break formation. *Mol. Cell. Biol.* **22**, 1106–1115 (2002).
47. Murakami, H., Borde, V., Nicolas, A. & Keeney, S. Gel electrophoresis assays for analyzing DNA double-strand breaks in *Saccharomyces cerevisiae* at various spatial resolutions. *Methods Mol. Biol.* **557**, 117–142 (2009).
48. Neale, M. J. & Keeney, S. End-labeling and analysis of Spo11-oligonucleotide complexes in *Saccharomyces cerevisiae*. *Methods Mol. Biol.* **557**, 183–195 (2009).
49. R Development Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2012).
50. Rumble, S. M. *et al.* SHRIMP: accurate mapping of short color-space reads. *PLOS Comput. Biol.* **5**, e1000386 (2009).
51. Abdullah, M. F. & Borts, R. H. Meiotic recombination frequencies are affected by nutritional states in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **98**, 14524–14529 (2001).
52. Cotton, V. E., Hoffmann, E. R., Abdullah, M. F. & Borts, R. H. Interaction of genetic and environmental factors in *Saccharomyces cerevisiae* meiosis: the devil is in the details. *Methods Mol. Biol.* **557**, 3–20 (2009).
53. Martini, E., Diaz, R. L., Hunter, N. & Keeney, S. Crossover homeostasis in yeast meiosis. *Cell* **126**, 285–295 (2006).
54. Vader, G. *et al.* Protection of repetitive DNA borders from self-induced meiotic instability. *Nature* **477**, 115–119 (2011).



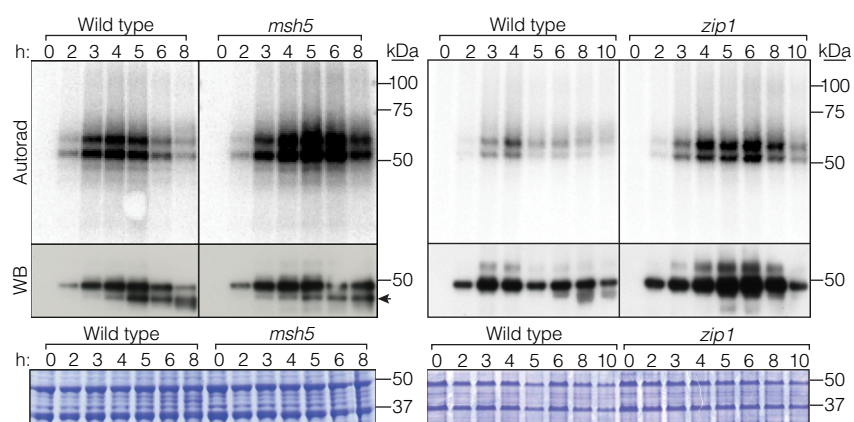
**Extended Data Figure 1 | Chromosomal breaks in *msh5* and *zip1* mutants.** Representative pulsed-field gel Southern blots probed for chromosome IX are shown, labelled as in Fig. 1b.



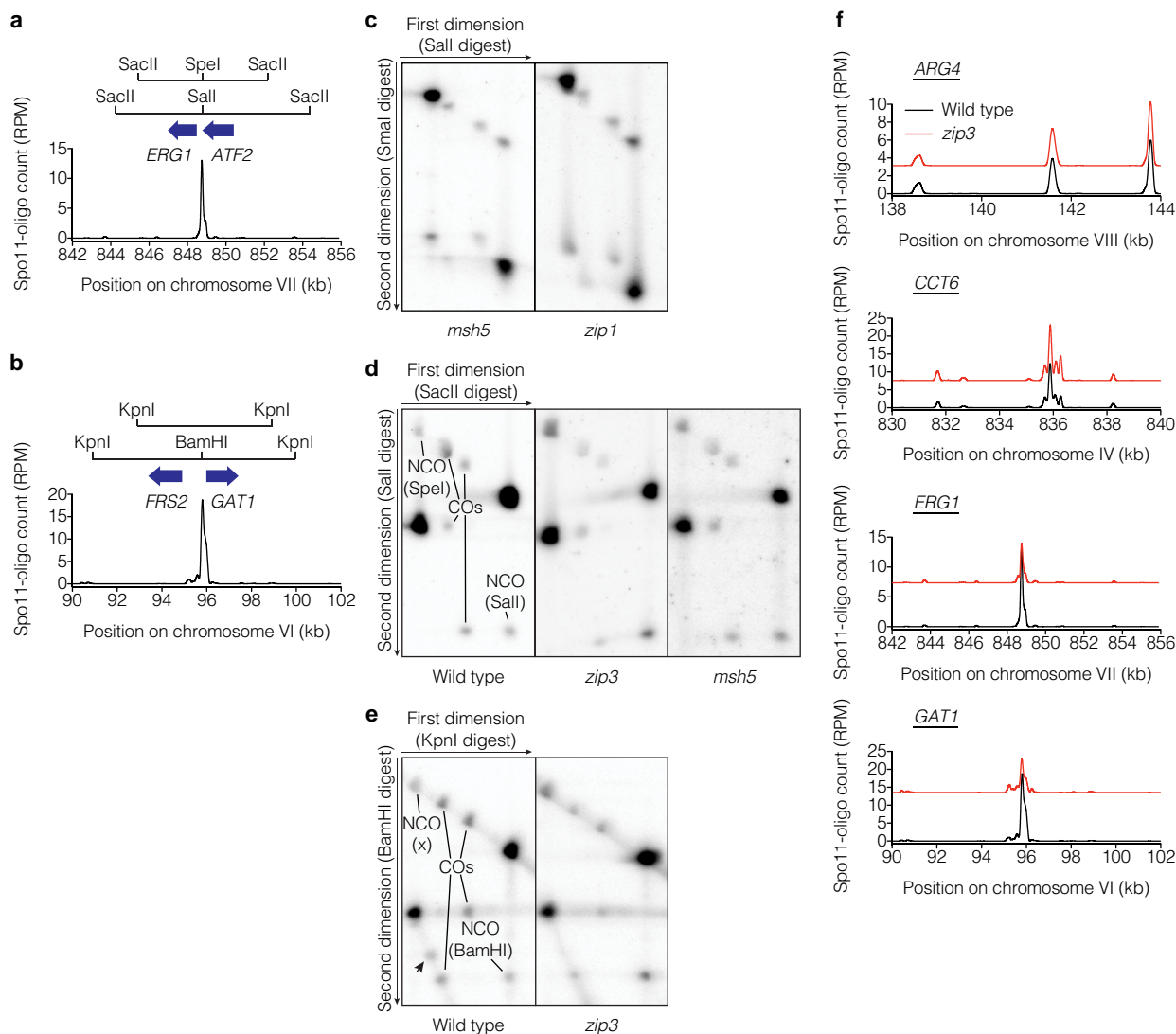
**Extended Data Figure 2 | DSB formation appears normal in *SPO11-Flag* and *SPO11-PrA* strains.** **a**, Southern blots probed for chromosome III. High molecular weight chromosomal DNA was purified 6 h after transfer to sporulation medium from meiotic *rad50S* cultures carrying the indicated *SPO11* alleles (in *spo11-yf* the catalytic tyrosine 135 is mutated to phenylalanine), then separated on pulsed-field electrophoresis gels. Samples from a *rad50S spo11-HA* strain are shown for comparison; haemagglutinin-tagged Spo11 has reduced DSB frequency. Each lane represents an independent

culture (*SPO11*<sup>+</sup> samples from the same cultures were run on both gels). PrA, protein A. **b**, Quantification of blots in panel **a** and separate blots (not shown) probed for chromosomes VII or VIII. Break frequencies are per cent of DNA in lane (mean ± s.d. of 3–4 cultures). Numbers in parentheses indicate values from each tagged strain relative to *SPO11*<sup>+</sup> for the same chromosome. Relative DSB frequencies at the bottom are averages across the three chromosomes assayed.



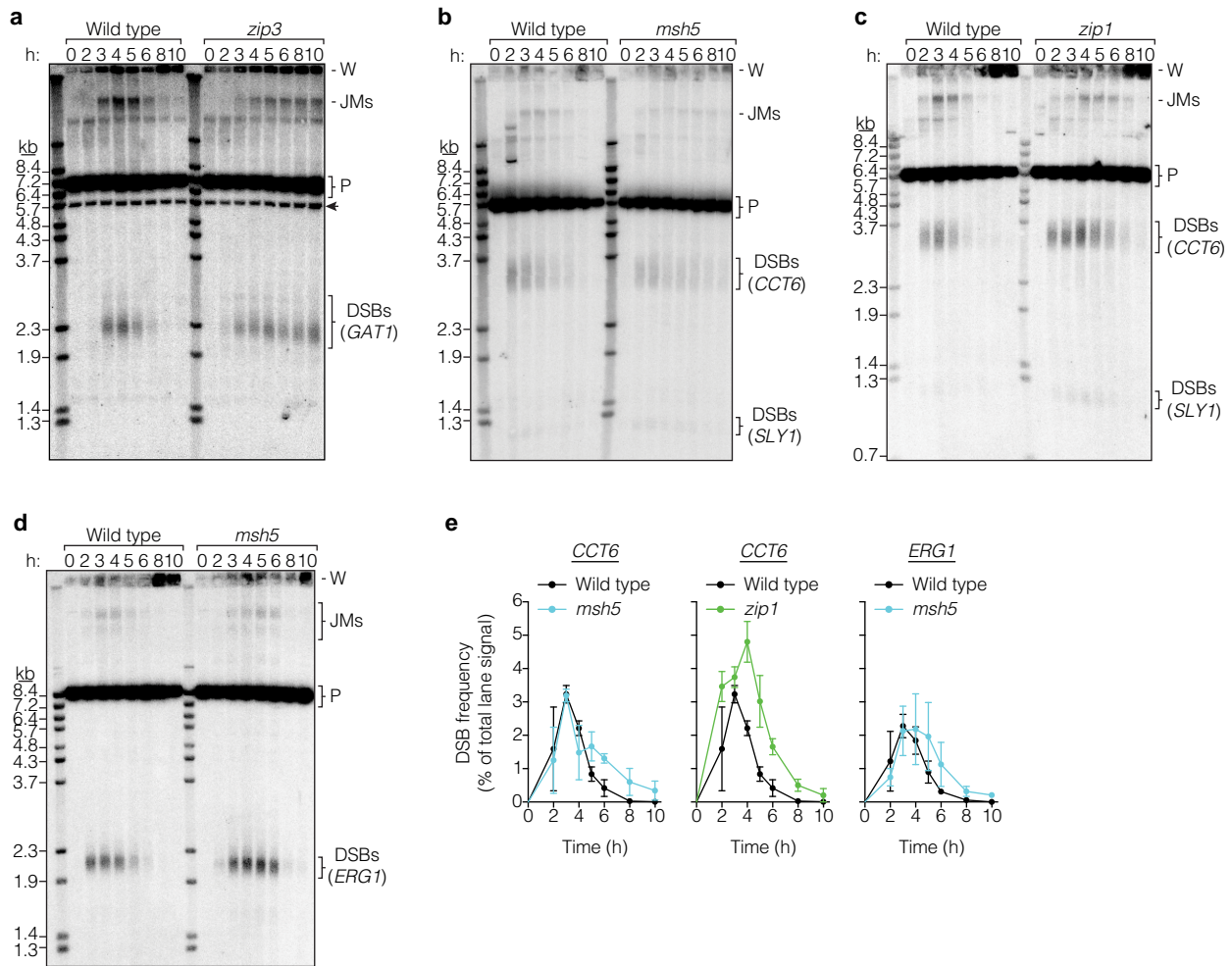


Extended Data Figure 3 | Spo11-oligo complexes in *msh5* and *zip1* mutants. Representative time courses are shown.



**Extended Data Figure 4 | Analysis of recombination at three natural DSB hotspots.** **a, b**, Recombination reporters at the *ERG1* (**a**) and *GAT1* (**b**) hotspots. **c–e**, Representative Southern blots of parental and recombinant DNA molecules at *CCT6* (**c**), *ERG1* (**d**) and *GAT1* (**e**). The arrowhead in

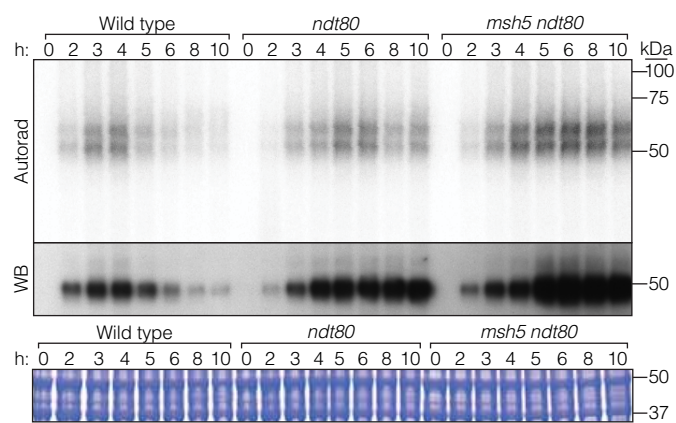
**e** indicates a non-reproducible radiolabelled species. **f**, Local distribution of DSBs around recombination reporter locations is not altered in *zip3* mutants. Spo11-oligo profiles (averages for wild type and *zip3* mutant) are smoothed with 201-bp Hann window; *zip3* values are offset to separate profiles.



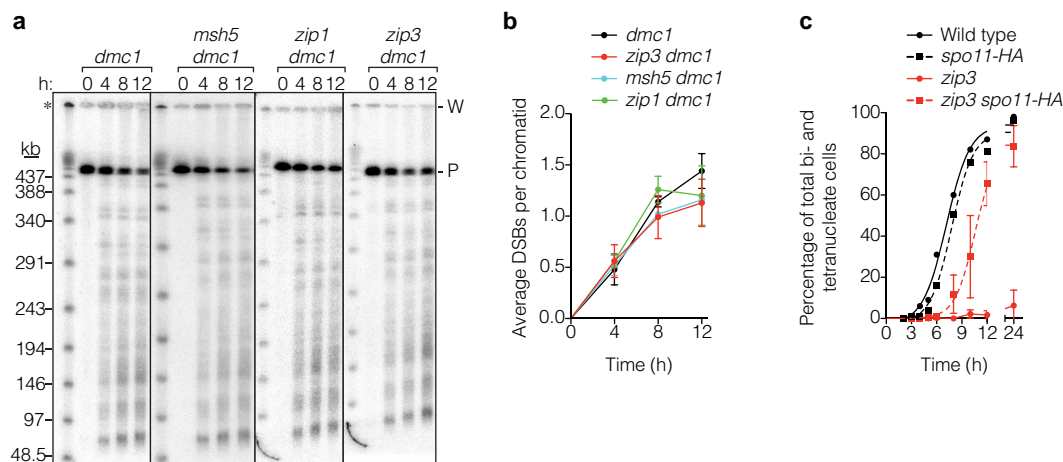
**Extended Data Figure 5 | Direct analysis of DSB formation at natural hotspots.** **a–d**, Representative Southern blots of DNA separated on a conventional agarose gel and probed for *GAT1* (**a**), *CCT6* (**b**, **c**) and *ERG1* (**d**).

The arrowhead in **a** indicates signal from the *CCT6* parental band that remained after stripping and reprobing for *GAT1*. **e**, Quantifications for **b–d** (mean  $\pm$  s.d. for 3 cultures).



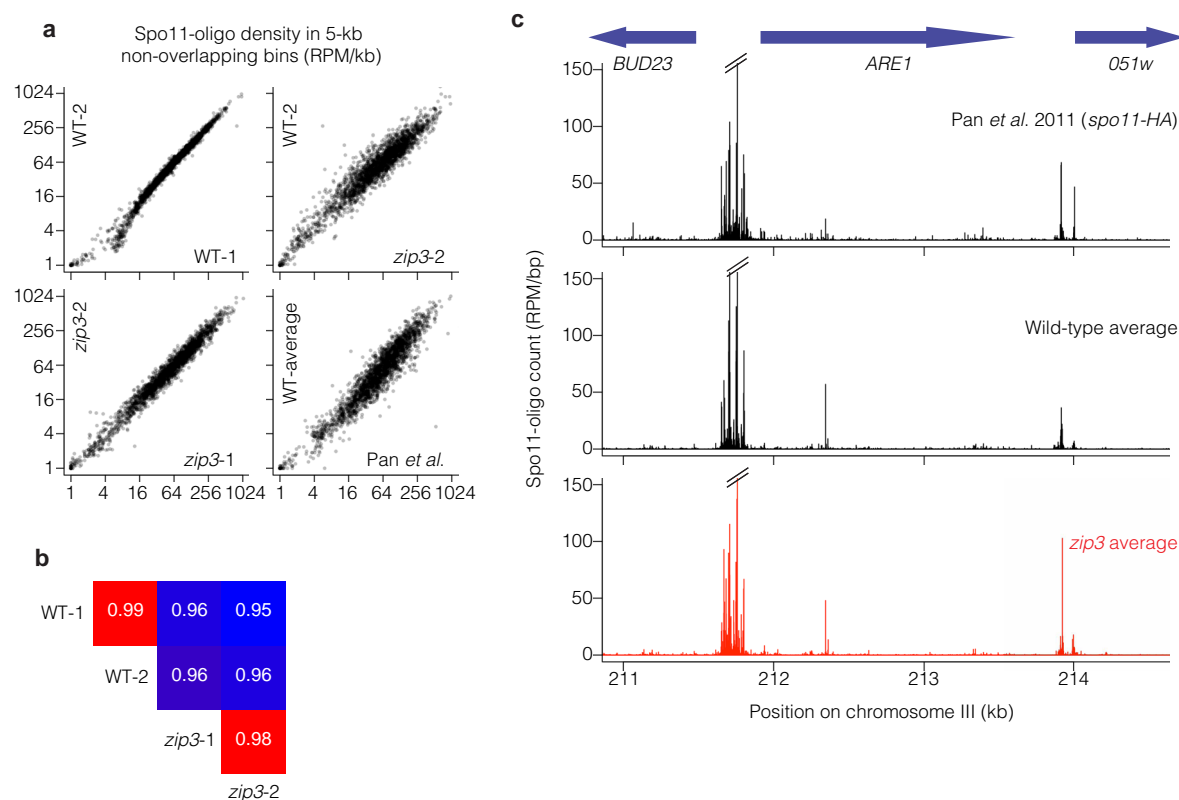


**Extended Data Figure 6 | Spo11-oligo complexes in *msh5 ndt80* double mutant.** Representative time courses are shown.



**Extended Data Figure 7 | Effects of *dmc1* deletion or *spo11* hypomorphic mutation on ZMM mutant phenotypes.** **a**, **b**, ZMM status is irrelevant in a *dmc1* background. Broken chromosomes accumulate to similar levels in a *dmc1* single mutant and *dmc1 zmm* double mutants. Representative pulsed-field gel Southern blots probed for chromosome IX are in **a** and Poisson-corrected quantification of DSBs is in **b** (mean  $\pm$  s.d., 3 cultures). **c**, Reducing Spo11

activity in a *zip3* mutant partially alleviates the prophase I delay/arrest. Meiotic progression was assessed by staining with DAPI (4',6-diamidino-2-phenylindole) and measuring the percentage of cells that had completed meiosis I (MI) with or without completing meiosis II ( $\pm$  MII). Data are means  $\pm$  s.d. for 3 cultures, except wild type and *spo11-HA*, each analysed once.

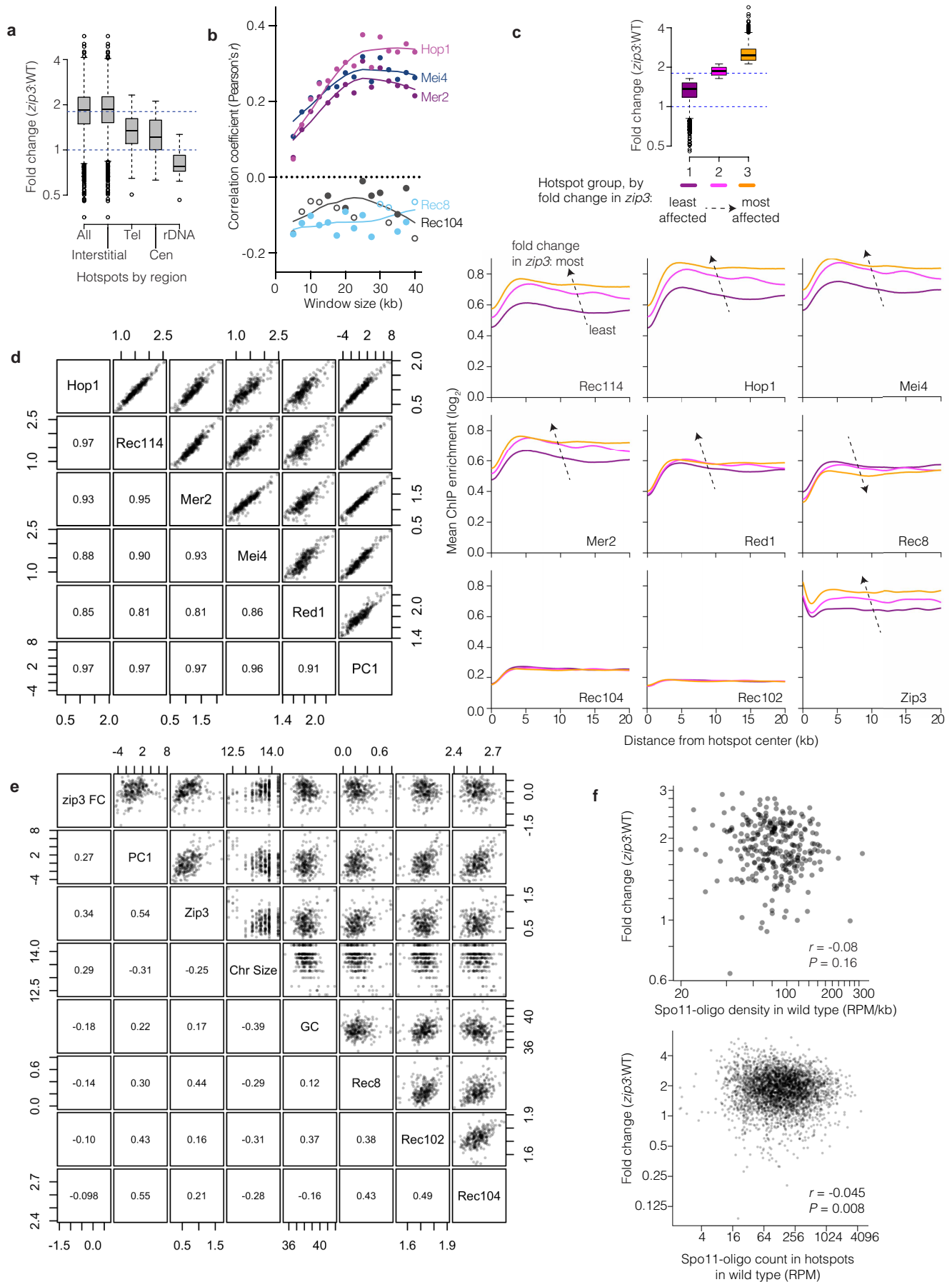


### Extended Data Figure 8 | Spo11-oligo mapping in wild type and *zip3* mutant.

**a, b**, Quantitative reproducibility of Spo11-oligo maps. In **a**, comparisons are shown for individual wild type (WT) or *zip3* data sets from the present study, or the previously published *spo11-HA* data (from ref. 11). Uniquely mapped Spo11 oligos were summed in non-overlapping 5-kb bins and expressed as RPM per kb (plotted on a log scale). In **b**, pairwise correlation coefficients for the data sets from the current study are shown (Pearson's  $r$ ; box colours scaled from blue to red proportional to strength of correlation). For the comparison of this study's wild-type average with data from Pan *et al.*,

$r = 0.949$ . Note that Pan *et al.* used a different strain background with different auxotrophies, which may alter DSB distributions<sup>51,52</sup>, and a hypomorphic *spo11* allele (*spo11-HA*), which may affect DSBs to different extents at different locations<sup>53</sup>. Note that biological replicates (WT-1 versus WT-2 or *zip3-1* versus *zip3-2*) agreed better than comparisons between cultures of different genotype. **c**, DSBs form at the same hotspots and with similar distribution within and between hotspots in wild type and *zip3*. Unsmoothed Spo11-oligo maps are shown in the vicinity of the well-characterized *ARE1* (*YCR048w*) hotspot.





**Extended Data Figure 9 | Changes in the DSB landscape in *zip3* mutant.**

**a**, Change in Spo11-oligo counts in hotspots grouped by chromosomal context. Tel, within 20 kb of telomeres; Cen, within  $\pm 10$  kb of centromeres; rDNA, from 60 kb leftward to 30 kb rightward of rDNA; Interstitial, all others. Dashed lines mark values assumed as no change and average change (1.8-fold). Boxes indicate median and interquartile range; whiskers indicate the most extreme data points which are  $\leq 1.5$  times the interquartile range from the box; individual points are outliers. Subtelomeric and pericentric zones show less increase in *zip3* on average, thus, ZMM-dependent feedback contributes less than other, unknown factors to suppressing DSBs in these regions. The zone near the rDNA showed no increase or was even decreased; thus, *zip3* mutants are competent for this region's DSB suppression, which is dependent on the ATPase Pch2 and the replication factor Orc1 (ref. 54). Note that the remaining interstitial hotspots showed highly variable response to *zip3* mutation ( $>20$  fold). **b**, Correlation between log-fold change in Spo11-oligo counts in *zip3* and the binding of the indicated proteins, binned in non-overlapping windows of varying size. Closed symbols,  $P < 0.05$ . ChIP data are from ref. 38. **c**, Average ChIP profiles around interstitial hotspots divided into three equal-sized groups according to the average fold change in *zip3*. Top, the box and whisker plot (as described for **a**) shows the distribution of fold changes for the three groups. Bottom, ChIP profiles for each of the indicated proteins. Note that the profiles lie atop one another for Rec102 and Rec104. Dashed arrows indicate direction

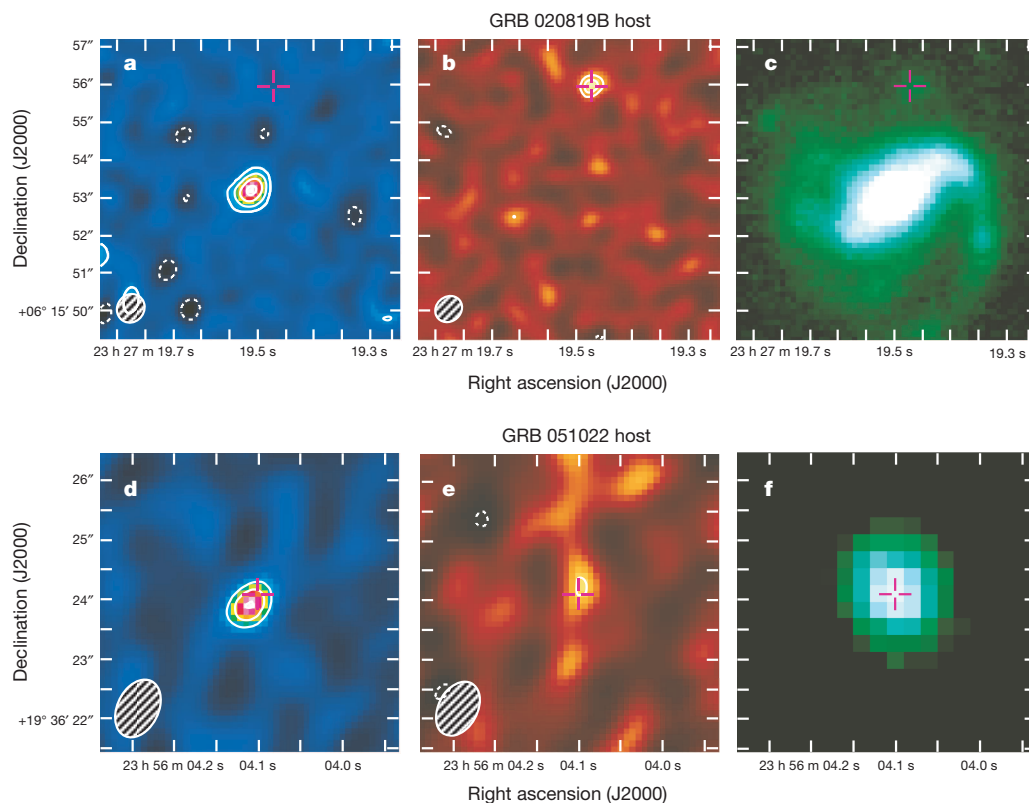
of the change in the average profiles with increasing fold change in *zip3*. ChIP data are from refs 38 and 39. **d**, High degree of colinearity of  $\log_2$ -transformed ChIP data<sup>38</sup> for Rec114, Mei4 and Mer2 (which are essential for DSB formation) and Hop1 and Red1 (axis proteins that promote normal DSB formation). More than 90% of the variance for this combination of ChIP data is captured in the first principal component (PC1). The high degree of correlation between these proteins was described previously<sup>38</sup>. **e**, Correlations between the fold change in *zip3* (*zip3* FC,  $\log_2$  and assuming 1.8-fold increase genome-wide) and various chromosomal features: principal component 1 for Rec114, Mei4, Mer2, Hop1 and Red1 ChIP data (same as in **d**); chromosome size ( $\log_e(\text{bp})$ ); G+C content (%); and ChIP data for the indicated proteins ( $\log_2$ ). In **d** and **e**, top right panels show pairwise scatter plots and bottom left panels show corresponding correlation coefficients (Pearson's  $r$ ) for data for interstitial regions binned in 35-kb non-overlapping windows. Essentially identical results were obtained with different window sizes (20–40 kb) or with varying placement of windows (data not shown). **f**, Essentially no correlation between DSB activity in wild type and change in *zip3*, whether considering interstitial regions divided into non-overlapping 35-kb bins (upper panel) or interstitial hotspots (lower panel). A 1.8-fold increase genome-wide in *zip3* is assumed. Note: fold change is labelled according to a linear scale but plotted in a log scale in panels **a**, **c**, **f**.

# Two $\gamma$ -ray bursts from dusty regions with little molecular gas

B. Hatsukade<sup>1</sup>, K. Ohta<sup>2</sup>, A. Endo<sup>3</sup>, K. Nakanishi<sup>1,4,5</sup>, Y. Tamura<sup>6</sup>, T. Hashimoto<sup>1</sup> & K. Kohno<sup>6,7</sup>

Long-duration  $\gamma$ -ray bursts are associated with the explosions of massive stars<sup>1</sup> and are accordingly expected to reside in star-forming regions with molecular gas (the fuel for star formation). Previous searches for carbon monoxide (CO), a tracer of molecular gas, in burst host galaxies did not detect any emission<sup>2–4</sup>. Molecules have been detected as absorption in the spectra of  $\gamma$ -ray burst afterglows, and the molecular gas is similar to the translucent or diffuse molecular clouds of the Milky Way<sup>5,6</sup>. Absorption lines probe the interstellar medium only along the line of sight, so it is not clear whether the molecular gas represents the general properties of the regions where the bursts occur. Here we report spatially resolved observations of CO line emission and millimetre-wavelength continuum emission in two galaxies hosting  $\gamma$ -ray bursts. The bursts happened in regions rich in dust, but not particularly rich in molecular gas. The ratio of molecular gas to dust ( $<9$ – $14$ ) is significantly lower than in star-forming regions of the Milky Way and nearby star-forming galaxies, suggesting that much of the dense gas where stars form has been dissipated by other massive stars.

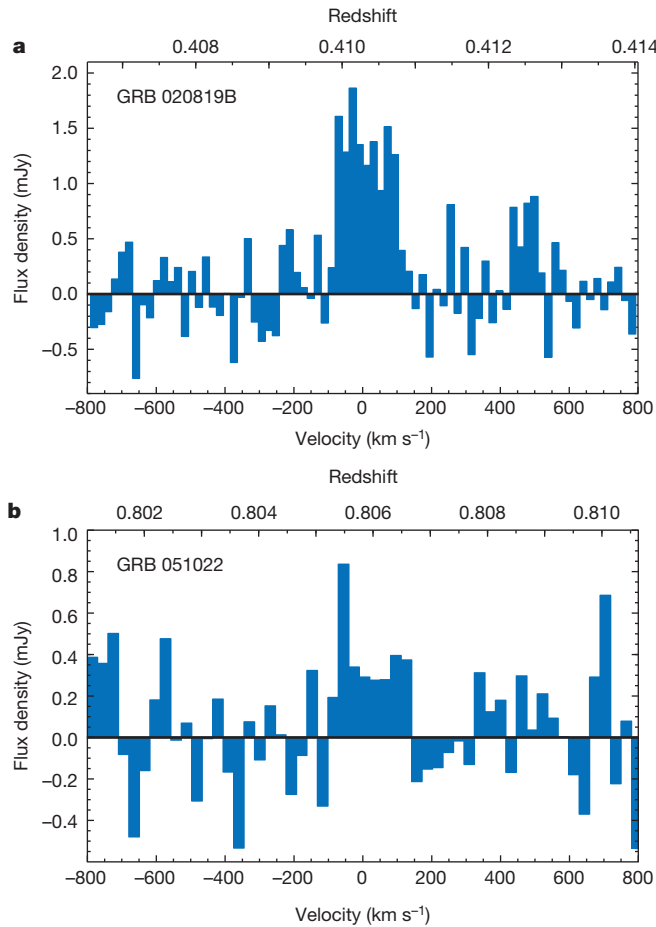
We selected the two  $\gamma$ -ray burst (GRB) hosts (GRB 020819B at a redshift of  $z = 0.41$  and GRB 051022 at  $z = 0.81$ ) with high star-formation rates (SFRs) and high gas metallicity among GRB host galaxies to maximize the possibility of detecting the CO emission line and dust continuum emission. The GRB 020819B host shows an extinction-corrected SFR of  $\sim 10$ – $30 M_{\odot} \text{ yr}^{-1}$  (where  $M_{\odot}$  indicates solar mass) derived from ultraviolet continuum emission, the H $\alpha$  emission line, and spectral energy distribution (SED) fitted using infrared data<sup>7–9</sup>. The SFRs at spatially resolved positions are also derived from the H $\alpha$  emission line, which are  $10.2 M_{\odot} \text{ yr}^{-1}$  and  $23.6 M_{\odot} \text{ yr}^{-1}$  at the nuclear region and at the GRB explosion site, respectively<sup>8</sup>. The host galaxy of GRB 051022 shows an extinction-corrected SFR of  $\sim 20$ – $70 M_{\odot} \text{ yr}^{-1}$  derived from ultraviolet continuum emission, the [O II] emission line at rest-frame wavelength  $\lambda = 3727 \text{ \AA}$ , the SED fitted with infrared data, and radio continuum emission<sup>7,9–11</sup>. The gas metallicity is measured at the GRB 020819B site, the nuclear region of the GRB 020819B host, and at the GRB 051022 host, and they all have at least solar metallicity<sup>8,12</sup>. The two GRBs are classified as ‘dark GRBs’<sup>13,14</sup>, whose afterglow



**Figure 1 | CO maps, 1.2-mm continuum maps and optical images of the GRB hosts.** The magenta cross represents the position of the radio afterglow. The ALMA beam size is shown in the lower left corners of **a**, **b**, **d** and **e** (black and white ovals). **a**, Velocity-integrated CO(3–2) intensity map. Contours start from  $\pm 3\sigma$  with  $2\sigma$  step ( $1\sigma = 0.040 \text{ Jy beam}^{-1} \text{ km s}^{-1}$ ). **b**, 1.2-mm continuum map. Contours start from  $\pm 3\sigma$  with  $1\sigma$  step ( $1\sigma = 0.030 \text{ mJy beam}^{-1}$ ). **c**, Optical R-band image obtained with the Gemini North Telescope. **d**, Velocity-integrated CO(4–3) intensity map. Contours start from  $3\sigma$  with  $1\sigma$  step ( $1\sigma = 0.037 \text{ Jy beam}^{-1} \text{ km s}^{-1}$ ). **e**, 1.2-mm continuum. Contours are  $\pm 3\sigma$  ( $1\sigma = 0.032 \text{ mJy beam}^{-1}$ ). **f**, Optical R-band image obtained with the Gemini South Telescope.

<sup>1</sup>National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan. <sup>2</sup>Department of Astronomy, Kyoto University, Kyoto 606-8502, Japan. <sup>3</sup>Kavli Institute of NanoScience, Faculty of Applied Sciences, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands. <sup>4</sup>Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago 763 0355, Chile. <sup>5</sup>The Graduate University for Advanced Studies (SOKENDAI), 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan. <sup>6</sup>Institute of Astronomy, University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan. <sup>7</sup>Research Centre for the Early Universe, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan.





**Figure 2 | CO spectra of the GRB hosts.** Continuum emission is subtracted. **a**, CO(3–2) spectrum of the nuclear region of the GRB 020819B host at  $20 \text{ km s}^{-1}$  resolution. A Gaussian fit to the emission line gives a redshift of  $z = 0.410$  and a velocity width of  $167 \text{ km s}^{-1}$  (FWHM). **b**, CO(4–3) spectrum of the GRB 051022 host at  $30 \text{ km s}^{-1}$  resolution. A Gaussian fit to the emission line gives a redshift of  $z = 0.806$  and a velocity width of  $176 \text{ km s}^{-1}$  (FWHM).

is optically dark compared with what is expected from X-ray afterglows<sup>15</sup>. The origin of the dark GRBs is not yet well understood, but a plausible explanation is dust obscuration along the line of sight to GRBs<sup>16–19</sup>.

We conducted observations of the CO emission line and 1.2-mm continuum towards the GRB hosts using the Atacama Large Millimeter/submillimeter Array (ALMA). We observed the redshifted CO(3–2) line for the GRB 020819B host and the CO(4–3) line for the GRB 051022 host. The angular resolution is  $\sim 0.8'' \times 0.7''$  ( $\sim 4 \text{ kpc}$

$\times 4 \text{ kpc}$ ) and  $\sim 1.0'' \times 0.7''$  ( $\sim 8 \text{ kpc} \times 5 \text{ kpc}$ ) (full-width at half maximum; FWHM) for the GRB 020819B host and the GRB 051022 host, respectively. The CO emission line is clearly detected at the nuclear region of the GRB 020819B host and the GRB 051022 host (Figs 1a and d and 2). While molecular gas has been detected in absorption in spectra of GRB afterglows<sup>5,6</sup>, this is the first case for detecting spatially resolved molecular gas emission in GRB hosts<sup>2–4</sup>. The component size of the CO emission (deconvolved from beam) derived from a Gaussian fitting is  $3.2 \times 1.5 \text{ kpc}$  (FWHM). The molecular gas mass estimated from the CO emission is  $M_{\text{gas}} = 2.4 \times 10^9 M_{\odot}$  and  $2.1 \times 10^9 M_{\odot}$  for the nuclear region of the GRB 020819B host and the GRB 051022 host, respectively (see Methods and Table 1). The molecular gas mass is comparable to those of local massive spiral galaxies<sup>20</sup>, and lower than those of  $z \approx 1$ –2 normal star-forming galaxies<sup>21</sup> or submillimetre-luminous galaxies<sup>22,23</sup>. The fraction of molecular gas mass to stellar mass<sup>7,24</sup> for the hosts is  $\sim 0.1$ , which is comparable to those of local spiral galaxies<sup>20</sup>.

The 1.2-mm continuum emission is also detected in both GRB hosts (Fig. 1b, e). The spatially resolved continuum map of the GRB 020819B host shows that the emission is significantly detected only at a star-forming region  $\sim 3''$  (16 kpc in projection) away from the nuclear region, where the GRB explosion occurred. The size of the continuum emission deconvolved from the telescope beam is  $\sim 1.7 \text{ kpc} \times 1.0 \text{ kpc}$ . We regard the continuum emission as dust thermal emission originating in star-forming activity (see Supplementary Information). By assuming that the dust emission is described as a modified blackbody and using the dust temperature and emissivity index derived from fitting, we derive the dust masses of  $M_{\text{dust}} = 4.8 \times 10^7 M_{\odot}$  and  $2.9 \times 10^7 M_{\odot}$  for the GRB 020819B site and the GRB 051022 host, respectively (see Methods). The far-infrared luminosity and SFR are  $L_{\text{FIR}} = 1.1 \times 10^{11} L_{\odot}$  (where  $L_{\odot}$  is the solar luminosity) and  $\text{SFR} = 18 M_{\odot} \text{ yr}^{-1}$  for the GRB 020819B site, and  $L_{\text{FIR}} = 1.9 \times 10^{11} L_{\odot}$  and  $\text{SFR} = 32 M_{\odot} \text{ yr}^{-1}$  for the GRB 051022 host, respectively. The SFRs are comparable to the extinction-corrected SFRs derived from ultraviolet and optical observations, suggesting that there is no sign of an extra, optically completely invisible portion of star formation that cannot be recovered by extinction correction.

Of particular interest is that the spatial distributions of molecular gas and dust are clearly different in the GRB 020819B host. The ratio of molecular gas mass to dust mass of the GRB 020819B host is  $>51$ –60 and  $<9$ –14 ( $3\sigma$  limits with uncertainty from dust mass) at the nuclear region and the GRB site, respectively. The ratio in the GRB site is significantly lower than that of the nuclear region, indicating that the GRB occurred under particular circumstances within the host. The molecular gas-to-dust ratio at the GRB site is also lower than those of the Milky Way and nearby star-forming galaxies<sup>25</sup>, suggesting that the star-forming environment where GRBs occur is different from those in local galaxies. While the correlation between gas-to-dust ratio

**Table 1 | Properties of GRB host galaxies**

Property	GRB 020819B nuclear region	GRB 020819B site	GRB 051022
$z_{\text{CO}}$	0.410		0.806
CO transition	3–2		4–3
$L'_{\text{CO}(1-0)}$ ( $\text{K km s}^{-1} \text{ pc}^2$ )	$(5.5 \pm 0.4) \times 10^8$	$<1.3 \times 10^8$	$(4.9 \pm 0.9) \times 10^9$
$M_{\text{gas}} (M_{\odot})$	$(2.4 \pm 0.2) \times 10^9$	$<5.4 \times 10^8$	$(2.1 \pm 0.4) \times 10^9$
$S_{1.2 \text{ mm}}$ (mJy)	$<0.12$	$0.14 \pm 0.03$	$0.10 \pm 0.03$
$M_{\text{dust}} (M_{\odot})$	$<4.2 \times 10^7$	$(4.8 \pm 1.0) \times 10^7$	$(2.9 \pm 0.9) \times 10^7$
$L_{\text{FIR}} (L_{\odot})$	$<9.3 \times 10^{10}$	$(1.1 \pm 0.2) \times 10^{11}$	$(1.9 \pm 0.6) \times 10^{11}$
$\text{SFR} (M_{\odot} \text{ yr}^{-1})$	$<16$	$18 \pm 4$	$32 \pm 10$
$M_{\text{gas}}/M_{\text{dust}}$	$>51$ –60	$<9$ –14	58–86

The errors represent root-mean-square ( $1\sigma$ ) uncertainties from the photometry error. The limits are  $3\sigma$ . We adopt a cosmology with  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.27$ , and  $\Omega_{\Lambda} = 0.73$ . For details, see Methods.  $z_{\text{CO}}$  is the redshift determined from the CO line.  $L'_{\text{CO}(1-0)}$  is the CO(1–0) luminosity derived from  $L'_{\text{CO}} = 3.25 \times 10^7 S_{\text{CO}} \Delta \nu_{\text{obs}}^{-2} D_L^2 (1+z)^{-3}$ , where  $L'_{\text{CO}}$  is in units of  $\text{K km s}^{-1} \text{ pc}^2$ ,  $S_{\text{CO}} \Delta \nu$  is the velocity-integrated flux in  $\text{Jy km s}^{-1}$ ,  $\nu_{\text{obs}}$  is the observed line frequency in GHz and  $D_L$  is the luminosity distance in Mpc. We assume a CO line ratio of CO(3–2)/CO(1–0) = 0.93 and CO(4–3)/CO(1–0) = 0.85, which are the values for the local star-forming galaxy M82.  $M_{\text{gas}}$  is the molecular gas mass derived from  $M_{\text{gas}} = x_{\text{CO}} L'_{\text{CO}(1-0)}$ , where  $x_{\text{CO}}$  is the CO-to-molecular gas mass conversion factor of a Galactic value  $x_{\text{CO}} = 4.3$  in units of  $M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$ .  $S_{1.2 \text{ mm}}$  is the 1.2-mm continuum flux.  $M_{\text{dust}}$  is the dust mass derived from  $M_{\text{dust}} = S_{\text{obs}} D_L^2 / [(1+z) \kappa_d(\nu_{\text{rest}}) B(\nu_{\text{rest}}, T_d)]$ , where  $S_{\text{obs}}$  is the observed flux density,  $\nu_{\text{rest}}$  is the rest frequency,  $\kappa_d(\nu_{\text{rest}})$  is the rest-frequency mass absorption coefficient and  $B(\nu_{\text{rest}}, T_d)$  is the Planck function.  $L_{\text{FIR}}$  is the far-infrared luminosity derived from  $L_{\text{FIR}} = 4\pi M_d \int_0^{\infty} \kappa_d(\nu) B(\nu, T_d) d\nu$ . SFR is the star-formation rate derived from  $\text{SFR} (\text{in } M_{\odot} \text{ yr}^{-1}) = 1.72 \times 10^{-10} L_{\text{FIR}} (\text{in } L_{\odot})$ .

and metallicity has been observed<sup>20,25</sup>, it is unlikely that the large difference between the GRB site and the nuclear region is attributable to the difference in metallicity because both regions have a similar metallicity<sup>8</sup>. The difference of distribution between molecular gas and dust is also seen in the GRB 051022 host and the GRB site seems to be a dust-rich region, although the angular resolution is not good enough to be certain. The possible reasons for the deficit of molecular gas in the GRB site are that much of the dense gas where stars form has been incorporated into stars, or dissipated by a strong interstellar ultraviolet radiation field, which is expected in regions with intense star formation. The lack of molecular gas in optical spectra of GRB afterglows has been reported<sup>26</sup> and a possible explanation is the dissociation of molecules by ambient ultraviolet radiation with 10–100 times the Galactic mean value from the star-forming regions where GRB progenitors reside<sup>26,27</sup>. GRB hosts with a mean ultraviolet radiation field of 35–350 times the Galactic mean value have been observed<sup>28</sup>. The molecular gas-to-dust ratio in GRB hosts could be an important indicator of an environment where GRBs occur.

The occurrence of GRB 020819B in a dust-rich region supports the idea that the dust extinction is the cause of the darkness of the optical afterglow<sup>13,24</sup>. The molecular gas-to-dust ratio in the GRB site is comparable to or lower than the ratios in nuclear regions of local galaxies of  $\sim 20$ –40 (ref. 29) and submillimetre-luminous galaxies of  $\sim 50$  (ref. 30), suggesting the existence of GRBs that could occur in dusty galaxies such as submillimetre-luminous galaxies.

## METHODS SUMMARY

We conducted ALMA observations of the GRB 020819B host and the GRB 051022 host at 245.072 GHz and 255.142 GHz, respectively, with a bandwidth of 1,875 MHz and with 24–27 antennas. The data were reduced with the Common Astronomy Software Applications package in a standard manner. The maps were processed with the CLEAN algorithm with Briggs weighting (with the *robust* parameter equal to 0.5). The final synthesized beam size (FWHM) is  $\sim 0.8'' \times 0.7''$  and  $\sim 1.0'' \times 0.7''$  for the GRB 020819B host and the GRB 051022 host, respectively. We derived the molecular gas mass of  $M_{\text{gas}} = (2.4 \pm 0.2) \times 10^9 M_{\odot}$  and  $(2.1 \pm 0.4) \times 10^9 M_{\odot}$  for the nuclear region of the GRB 020819B host and the GRB 051022 host, respectively. Here we assume CO line ratios of  $\text{CO}(3-2)/\text{CO}(1-0) = 0.93$  and  $\text{CO}(4-3)/\text{CO}(1-0) = 0.85$ , which are the values for the local star-forming galaxy M82, by considering the star-forming property of the hosts. We adopt the CO-to-molecular-gas mass conversion of Galactic value  $(4.3 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1})$  because the metallicity of the two hosts is close to the solar metallicity. To estimate a dust temperature and an emissivity index, we fitted a single temperature modified blackbody form to the far-infrared–millimetre photometry of the ALMA 1.2-mm data and the Herschel Space Observatory 100- $\mu\text{m}$ , 160- $\mu\text{m}$  and 250- $\mu\text{m}$  data. The best-fitting results are  $T_{\text{dust}} = 28 \pm 3 \text{ K}$  and  $\beta = 1.9 \pm 0.3$  for the GRB 020819B host and  $T_{\text{dust}} = 34 \pm 6 \text{ K}$  and  $\beta = 1.8 \pm 0.5$  for the GRB 051022 host. By using the best-fitting modified blackbody functions, we derived dust masses of  $(4.8 \pm 1.0) \times 10^7 M_{\odot}$  and  $(2.9 \pm 0.9) \times 10^7 M_{\odot}$  for the GRB 020819B site and the GRB 051022 host, respectively.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 August 2013; accepted 3 April 2014.

1. Stanek, K. *et al.* Spectroscopic discovery of the supernova 2003dh associated with GRB 030329. *Astrophys. J.* **591**, L17–L20 (2003).
2. Kohn, K. *et al.* Nobeyama Millimeter Array observations of GRB 030329: a decay of afterglow with bumps and molecular gas in the host galaxy. *Publ. Astron. Soc. Jpn* **57**, 147–153 (2005).
3. Endo, A. *et al.* A revised estimate of the CO  $J = 1-0$  emission from the host galaxy of GRB 030329 using the Nobeyama Millimeter Array. *Astrophys. J.* **659**, 1431–1437 (2007).
4. Hatsukade, B., Kohn, K., Endo, A., Nakanishi, K. & Ohta, K. CO observations of the host galaxy of GRB 000418 at  $z = 1.1$ . *Astrophys. J.* **738**, 33–36 (2011).
5. Prochaska, J. X. *et al.* The first positive detection of molecular gas in a GRB host galaxy. *Astrophys. J.* **691**, L27–L32 (2009).
6. Krühler, T. *et al.* Molecular hydrogen in the damped Lyman  $\alpha$  system towards GRB 120815A at  $z = 2.36$ . *Astron. Astrophys.* **557**, 18–38 (2013).
7. Savaglio, S., Glazebrook, K. & Le Borgne, D. The galaxy population hosting gamma-ray bursts. *Astrophys. J.* **691**, 182–211 (2009).

8. Levesque, E. M., Kewley, L. J., Graham, J. F. & Fruchter, A. S. A high-metallicity host environment for the long-duration GRB 020819B. *Astrophys. J.* **712**, L26–L30 (2010).
9. Hunt, L. K. *et al.* New light on gamma-ray burst host galaxies with Herschel. *Astron. Astrophys.* **565**, 112–130 (2014).
10. Castro-Tirado, A. J. *et al.* The dark nature of GRB 051022 and its host galaxy. *Astron. Astrophys.* **475**, 101–107 (2007).
11. Perley, D. A. & Perley, R. A. Radio constraints on heavily-obscured star-formation within dark gamma-ray burst host galaxies. *Astrophys. J.* **778**, 172–189 (2013).
12. Graham, J. F. *et al.* Unusually high metallicity host of the dark LGRB 051022. *AIP Conf. Ser.* **1133**, 269–271 (2009).
13. Jakobsson, P. *et al.* The radio afterglow and host galaxy of the dark GRB 020819B. *Astrophys. J.* **629**, 45–51 (2005).
14. Rol, E. *et al.* GRB 051022: physical parameters and extinction of a prototype dark burst. *Astrophys. J.* **669**, 1098–1106 (2007).
15. Jakobsson, P. *et al.* Swift identification of dark gamma-ray bursts. *Astrophys. J.* **617**, L21–L24 (2004).
16. Perley, D. A. *et al.* The host galaxies of Swift dark gamma-ray bursts: observational constraints on highly obscured and very high redshift GRBs. *Astron. J.* **138**, 1690–1708 (2009).
17. Cenko, S. B. *et al.* Dark bursts in the Swift era: the Palomar 60 Inch-Swift Early Optical Afterglow Catalog. *Astrophys. J.* **693**, 1484–1493 (2009).
18. Greiner, J. *et al.* The nature of ‘dark’ gamma-ray bursts. *Astron. Astrophys.* **526**, 30–39 (2011).
19. Campana, S. *et al.* The X-ray absorbing column density of a complete sample of bright Swift gamma-ray bursts. *Mon. Not. R. Astron. Soc.* **421**, 1697–1702 (2012).
20. Leroy, A. K. *et al.* Heracles: the HERA CO line extragalactic survey. *Astron. J.* **137**, 4670–4696 (2009).
21. Tacconi, L. J. *et al.* Phibss: Molecular gas content and scaling relations in  $z \sim 1$ –3 massive, main-sequence star-forming galaxies. *Astrophys. J.* **768**, 74–95 (2013).
22. Bothwell, M. S. *et al.* A survey of molecular gas in luminous sub-millimetre galaxies. *Mon. Not. R. Astron. Soc.* **429**, 3047–3067 (2013).
23. Solomon, P. M. & Vanden Bout, P. A. Molecular gas at high redshift. *Annu. Rev. Astron. Astrophys.* **43**, 677–725 (2005).
24. Küpcü Yoldaş, A., Greiner, J., Klose, S., Krühler, T. & Savaglio, S. Highly extinguished host galaxy of the dark GRB 020819B. *Astron. Astrophys.* **515**, L2–L5 (2010).
25. Draine, B. T. *et al.* Dust masses, PAH abundances, and starlight intensities in the SINGS galaxy sample. *Astrophys. J.* **663**, 866–894 (2007).
26. Tumlinson, J., Prochaska, J. X., Chen, H.-W., Dessauges-Zavadsky, M. & Bloom, J. S. Missing molecular hydrogen and the physical conditions of GRB host galaxies. *Astrophys. J.* **668**, 667–673 (2007).
27. Whalen, D., Prochaska, J. X., Heger, A. & Tumlinson, J. The molecular hydrogen deficit in gamma-ray burst afterglows. *Astrophys. J.* **682**, 1114–1123 (2008).
28. Chen, H.-W. *et al.* High-redshift starbursting dwarf galaxies revealed by  $\gamma$ -ray burst afterglows. *Astrophys. J.* **691**, 152–174 (2009).
29. Seaquist, E., Yao, L., Dunne, L. & Cameron, H. Revised masses of dust and gas of SCUBA Local Universe Survey far-infrared bright galaxies based on a recent CO survey. *Mon. Not. R. Astron. Soc.* **349**, 1428–1434 (2004).
30. Kovács, A. *et al.* SHARC-2 350  $\mu\text{m}$  observations of distant submillimeter-selected galaxies. *Astrophys. J.* **650**, 592–603 (2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We acknowledge A. Kawamura and ALMA staff members for support. We thank K. Yabe and A. Seko for discussions. B.H. was supported by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for JSPS Fellows. K.O. was supported by a JSPS Grant-in-Aid for Scientific Research (C) (grant number 24540230). A.E. was supported by the NWO (Veni grant number 639.041.023). Y.T. was supported by JSPS Grant-in-Aid for Scientific Research on Innovative Areas (grant number 25103503). ALMA is a partnership of ESO (representing its member states), the NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. This research is based in part on observations made with Herschel. Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. This research makes use of data based on observations obtained at the Gemini Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (USA), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), the Ministério da Ciência, Tecnologia e Inovação (Brazil) and the Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina).

**Author Contributions** B.H. led the project, reduced the ALMA data and wrote the manuscript. K.O. conducted the photometry of the Gemini and Herschel data. All authors contributed to the ALMA proposal, discussed the results and implications, and commented on the manuscript.

**Author Information** This research is based on the following ALMA data: ADS/JAO.ALMA#2011.0.00232.S (available from the ALMA archive at <http://almascience.nao.ac.jp/almadata/archive>). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.H. (bunyo.hatsukade@nao.ac.jp).

## METHODS

**Observations, data reduction and results.** We conducted ALMA band-6 observations of the GRB 020819B host on 2012 November 17 with 27 antennas and the GRB 051022 host on 2012 November 21 and December 2 with 24 antennas during the ALMA cycle-0 session. The range of baseline lengths of the configuration is 15–402 m and 15–382 m for the observations of the GRB 020819B host and the GRB 051022 host, respectively. The maximum recoverable scale (the largest angular structure to which a given array is sensitive) for the array configurations is  $10''$ , which is large enough to cover the angular scale of the host galaxies. The correlator was used in the frequency domain mode with a bandwidth of 1,875 MHz ( $488.28 \text{ kHz} \times 3,840 \text{ channels}$ ). Four basebands were used, giving a total bandwidth of 7.5 GHz. We observed the redshifted CO(3–2) line at 245.072 GHz for the GRB 020819B host and the redshifted CO(4–3) line at 255.142 GHz for the GRB 051022 host. Uranus was observed as a flux calibrator and a quasar J2253 + 161 (3C454.3) was observed for bandpass and phase calibrations. The on-source time is 47 min and 71 min for the GRB 020819B host and the GRB 051022 host, respectively. The data were reduced with the Common Astronomy Software Applications<sup>31</sup> package in a standard manner. The maps were processed with the CLEAN algorithm with the Briggs weighting (with *robust* parameter of 0.5). The final synthesized beam size (FWHM) is  $\sim 0.8'' \times 0.7''$  and  $\sim 1.0'' \times 0.7''$  for the GRB 020819B host and the GRB 051022 host, respectively. 1.2-mm continuum maps were created with a total bandwidth of  $\sim 7.5 \text{ GHz}$ , excluding channels with emission line. CO emission and 1.2-mm continuum emission are detected at both GRB host galaxies (Figs 1a, b, d and e, and 2). The GRB 020819B host is spatially resolved in the observations, while the GRB 051022 host is not. The velocity-integrated CO intensity is  $S_{\text{CO}(3-2)} = 0.53 \pm 0.04 \text{ Jy km s}^{-1}$  and  $S_{\text{CO}(4-3)} = 0.19 \pm 0.03 \text{ Jy km s}^{-1}$  at the nucleus of the GRB 020819B host and the GRB 051022 host, respectively. The 1.2-mm continuum flux density is  $S_{1.2 \text{ mm}} = 0.14 \pm 0.03 \text{ mJy}$  and  $S_{1.2 \text{ mm}} = 0.10 \pm 0.03 \text{ mJy}$  at the explosion site of GRB 020819B and the GRB 051022 host, respectively.

**Molecular gas mass.** CO luminosity is derived from  $L'_{\text{CO}} = 3.25 \times 10^7 S_{\text{CO}} \Delta v_{\text{obs}}^{-2} D_L^2 (1+z)^{-3}$  (ref. 23), where  $L'_{\text{CO}}$  is in units of  $\text{K km s}^{-1} \text{ pc}^2$ ,  $S_{\text{CO}} \Delta v$  is the velocity-integrated flux in  $\text{Jy km s}^{-1}$ ,  $v_{\text{obs}}$  is the observed line frequency in GHz, and  $D_L$  is the luminosity distance in Mpc. We assume a CO line ratio of  $\text{CO}(3-2)/\text{CO}(1-0) = 0.93$  and  $\text{CO}(4-3)/\text{CO}(1-0) = 0.85$ , which are the values for the local star-forming galaxy M82 (ref. 32), by considering the star-forming property of the host galaxies. The derived CO(1–0) luminosities are  $(5.5 \pm 0.4) \times 10^8 \text{ K km s}^{-1} \text{ pc}^2$  and  $(4.9 \pm 0.9) \times 10^9 \text{ K km s}^{-1} \text{ pc}^2$  for the nuclear region of the GRB 020819B host and the GRB 051022 host, respectively. The molecular gas mass is derived from  $M_{\text{gas}} = \alpha_{\text{CO}} L'_{\text{CO}(1-0)}$ , where  $\alpha_{\text{CO}}$  is the CO-to-molecular-gas mass conversion factor in units of  $M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$  including He mass. It is thought that there is a correlation between  $\alpha_{\text{CO}}$  and metallicity in the local Universe and at  $z \approx 1-2$  (refs 33, 34);  $\alpha_{\text{CO}}$  decreases with increasing metallicity. Because the metallicity of the two hosts is close to the solar metallicity, we adopt the Galactic value of  $\alpha_{\text{CO}} = 4.3 M_{\odot} (\text{K km s}^{-1} \text{ pc}^2)^{-1}$  (ref. 35). The derived molecular gas masses are  $M_{\text{gas}} = (2.4 \pm 0.2) \times 10^9 M_{\odot}$  and  $(2.1 \pm 0.4) \times 10^9 M_{\odot}$  for the nuclear region of the GRB 020819B host and the GRB 051022 host, respectively.

**Photometry of Herschel Space Observatory<sup>36</sup> data.** We used the Herschel Photodetector Array Camera and Spectrometer (PACS)<sup>37</sup> data in the archive. We conducted aperture photometry on the 160- $\mu\text{m}$  image of the GRB 051022 host with SExtractor<sup>38</sup> and obtained a flux density of  $S_{160 \mu\text{m}} = 12 \text{ mJy}$  (with about 30% photometry error). There is no significant contamination from nearby sources to the photometry. The FWHM of the source size is  $\sim 14''$  at 160  $\mu\text{m}$ , which is comparable to the FWHM of PACS beam size<sup>37</sup>. We also measured the centroid of the 100- $\mu\text{m}$  emission of the GRB 020819B host and found that the emission is in between the galaxy centre and the peak of 1.2-mm continuum. It is possible that dust is more widely spread in the host galaxy, although the angular resolution is inadequate (FWHM of  $\sim 7''$ ).

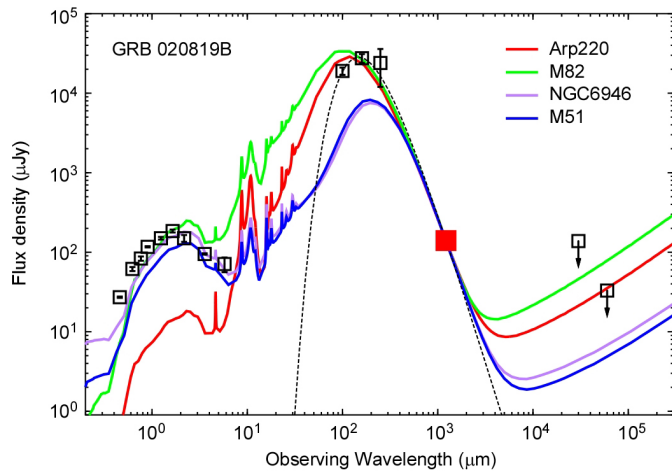
**Modified blackbody fit.** To estimate a dust temperature ( $T_{\text{dust}}$ ) and an emissivity index ( $\beta$ ), we fitted the far-infrared–millimetre photometry data of Herschel at 100  $\mu\text{m}$ , 160  $\mu\text{m}$  and 250  $\mu\text{m}$  (ref. 9), and ALMA at 1.2 mm with a single temperature modified blackbody form of  $S_{\nu} \propto \nu^{3+\beta}/(\exp(h\nu/kT_{\text{dust}}) - 1)$ , where  $S_{\nu}$  is the flux density and  $\nu$  is the frequency (Extended Data Fig. 1). The best-fitting results are  $T_{\text{dust}} = 28 \pm 3 \text{ K}$  and  $\beta = 1.9 \pm 0.3$  for the GRB 020819B host and  $T_{\text{dust}} = 34 \pm 6 \text{ K}$  and  $\beta = 1.8 \pm 0.5$  for the GRB 051022 host. We note that the missing flux in the ALMA observations in the scale of the PACS beam size

is negligible. The dust temperatures are within the typical range of  $z \approx 0-2$  star-forming galaxies<sup>39,40</sup>. The dust temperatures of the hosts were derived in a previous study with a SED model fit to optical-infrared data including Herschel photometry<sup>9</sup>:  $T_{\text{dust}} = 24.4 \text{ K}$  and  $52.6 \text{ K}$  for the GRB 020819B host and the GRB 051022 host, respectively. The dust temperature of the GRB 051022 host is higher than in this work. This may be due to their lack of photometric data at  $>160 \mu\text{m}$ , which is essential to fit dust SED.

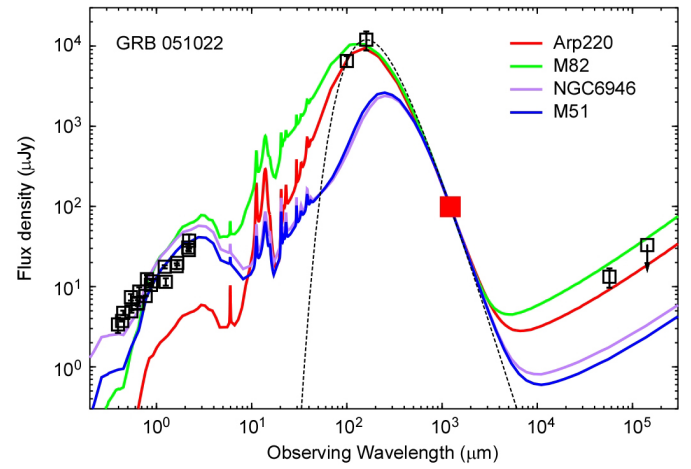
**Dust mass, far-infrared luminosity and SFR.** By using the best-fitting modified blackbody functions, we estimated dust mass, far-infrared luminosity and SFR. Dust mass is derived by  $M_{\text{dust}} = S_{\text{obs}} D_L^2 / [(1+z) \kappa_d(\nu_{\text{rest}}) B(\nu_{\text{rest}} T_{\text{dust}})]$  (ref. 41), where  $S_{\text{obs}}$  is the observed flux density,  $\nu_{\text{rest}}$  is the rest frequency,  $\kappa_d(\nu_{\text{rest}})$  is the rest-frequency mass absorption coefficient,  $B(\nu_{\text{rest}} T_{\text{dust}})$  is the Planck function. We assume that the absorption coefficient varies as  $\kappa_d(\nu) \propto \nu^{\beta}$  and  $\kappa_d(125 \mu\text{m}) = 26.4 \text{ cm}^2 \text{ g}^{-1}$  (ref. 42). The derived dust mass is  $(4.8 \pm 1.0) \times 10^7 M_{\odot}$  and  $(2.9 \pm 0.9) \times 10^7 M_{\odot}$  for the GRB 020819B site and the GRB 051022 host, respectively. If we use the dust temperature of 52.6 K for the GRB 051022 host estimated in a previous work<sup>9</sup>, the derived dust mass would be about a factor of two lower, which has no effect on the discussion in the main text. Far-infrared luminosity  $L_{\text{FIR}}$  is derived from  $L_{\text{FIR}} = 4\pi M_{\text{dust}} \int_0^{\infty} \kappa_d(\nu) B(\nu, T_{\text{dust}}) d\nu$  (ref. 41). The derived far-infrared luminosity is  $(1.1 \pm 0.2) \times 10^{11} L_{\odot}$  and  $(1.9 \pm 0.6) \times 10^{11} L_{\odot}$  for the GRB 020819B site and the GRB 051022 host, respectively. SFR is derived from the far-infrared luminosity as follows:  $\text{SFR (in } M_{\odot} \text{ yr}^{-1}) = 1.72 \times 10^{-10} L_{\text{FIR (in } L_{\odot})}$  (ref. 43), and calculated to be  $18 \pm 4 M_{\odot} \text{ yr}^{-1}$  and  $32 \pm 10 M_{\odot} \text{ yr}^{-1}$  for the GRB 020819B site and the GRB 051022 host, respectively. The comparison of CO and far-infrared luminosities is shown in Extended Data Fig. 2.

31. McMullin, J. P. *et al.* CASA Architecture and Applications. *Astronomical Data Analysis Software and Systems XVI* (eds Shaw, R. A., Hill, F. & Bell, D. J.) ASP Conf. Ser. 376 127–130 (ASP, 2007).
32. Carilli, C. L. & Walter, F. Cool gas in high-redshift galaxies. *Annu. Rev. Astron. Astrophys.* **51**, 105–161 (2013).
33. Wilson, C. D. The metallicity dependence of the CO-to-H<sub>2</sub> conversion factor from observations of local group galaxies. *Astrophys. J.* **448**, L97–L100 (1995).
34. Genzel, R. *et al.* The metallicity dependence of the CO  $\rightarrow$  H<sub>2</sub> conversion factor in  $z \geq 1$  star-forming galaxies. *Astrophys. J.* **746**, 69–79 (2012).
35. Bolatto, A. D., Wolfire, M. & Leroy, A. K. The CO-to-H<sub>2</sub> conversion factor. *Annu. Rev. Astron. Astrophys.* **51**, 207–268 (2013).
36. Pilbratt, G. L. *et al.* Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy. *Astron. Astrophys.* **518**, L1 (2010).
37. Poglitsch, A. *et al.* The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory. *Astron. Astrophys.* **518**, L2 (2010).
38. Bertin, E. & Arnouts, S. SExtractor: software for source extraction. *Astron. Astrophys.* **117** (Suppl.), 393–404 (1996).
39. Elbaz, D. *et al.* GOODS-Herschel: an infrared main sequence for star-forming galaxies. *Astron. Astrophys.* **533**, 119–144 (2011).
40. Symeonidis, M. *et al.* The Herschel census of infrared SEDs through cosmic time. *Mon. Not. R. Astron. Soc.* **431**, 2317–2340 (2013).
41. De Breuck, C. *et al.* CO emission and associated HI absorption from a massive gas reservoir surrounding the  $z = 3$  radio galaxy B3 J2330+3927. *Astron. Astrophys.* **401**, 911–925 (2003).
42. Dunne, L. *et al.* A census of metals at high and low redshift and the connection between submillimetre sources and spheroid formation. *Mon. Not. R. Astron. Soc.* **341**, 589–598 (2003).
43. Kennicutt, R. C., Jr. The global Schmidt law in star-forming galaxies. *Astrophys. J.* **498**, 541–552 (1998).
44. Svensson, K. M., Levan, A. J., Tanvir, N. R., Fruchter, A. S. & Strolger, L.-G. The host galaxies of core-collapse supernovae and gamma-ray bursts. *Mon. Not. R. Astron. Soc.* **405**, 57–76 (2010).
45. Stanway, E. R., Davies, L. J. M. & Levan, A. J. Low radio-derived star formation rates in  $z < 0.5$  gamma-ray burst host galaxies. *Mon. Not. R. Astron. Soc.* **409**, L74–L78 (2010).
46. Hatsukade, B. *et al.* Constraints on obscured star formation in host galaxies of gamma-ray bursts. *Astrophys. J.* **748**, 108–111 (2012).
47. Silva, L. *et al.* Modeling the effects of dust on galactic spectral energy distributions from the ultraviolet to the millimeter band. *Astrophys. J.* **509**, 103–117 (1998).
48. Gao, Y. & Solomon, P. M. The star formation rate and dense molecular gas in galaxies. *Astrophys. J.* **606**, 271–290 (2004).
49. Combes, F. *et al.* Galaxy evolution and star formation efficiency at  $0.2 < z < 0.6$ . *Astron. Astrophys.* **528**, 124–134 (2011).
50. Combes, F. *et al.* Gas fraction and star formation efficiency at  $z < 1.0$ . *Astron. Astrophys.* **550**, 41–55 (2013).
51. Daddi, E. *et al.* Very high gas fractions and extended gas reservoirs in  $z = 1.5$  disk galaxies. *Astrophys. J.* **713**, 686–707 (2010).

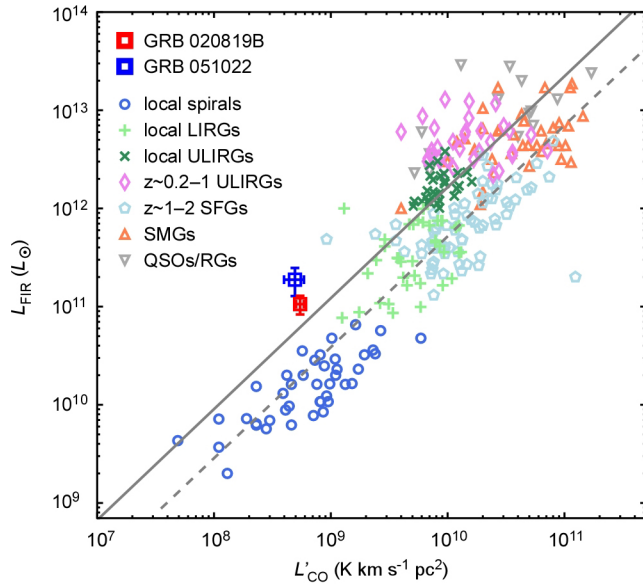




**Extended Data Figure 1 | Spectral energy distribution of the GRB 020819B host and the GRB 051022 host.** The red squares show ALMA 1.2-mm data. Black squares represent photometry from the literature<sup>10,11,14,24,44,45,46</sup> and the publicly archived data of Herschel. Dashed curves show the best-fit modified



blackbody functions. The arrows represent  $3\sigma$  upper limits. For comparison, we plot SED models of Arp220, M82, NGC6946 and M51 (ref. 47). The SED models are scaled to the flux density of ALMA data.



#### Extended Data Figure 2 | Comparison of CO and far-infrared luminosities.

The GRB 020819B host and the GRB 051022 host are plotted with  $1\sigma$  uncertainties (red and blue squares). To examine the properties of the GRB host galaxies as a whole and to compare with previous studies, we plot our data without separating the nuclear region and the explosion site for the GRB 020819B host galaxy. Various galaxy populations are also plotted: local spirals<sup>20,48</sup> (circles), local luminous infrared galaxies (LIRGs) (plus symbols) and ultraluminous infrared galaxies (ULIRGs)<sup>23,48</sup> (crosses),  $z \approx 0.2-1$  ULIRGs<sup>49,50</sup> (diamonds),  $z \approx 1-2$  normal star-forming galaxies<sup>21</sup> (pentagons), submillimetre-luminous galaxies<sup>22,23</sup> (up-triangles), QSOs and radio galaxies<sup>23</sup> (down-triangles). The grey solid and dashed lines represent the sequence of normal star-forming galaxies and starburst galaxies, respectively<sup>51</sup>.

# Nanotwinned diamond with unprecedented hardness and stability

Quan Huang<sup>1\*</sup>, Dongli Yu<sup>1\*</sup>, Bo Xu<sup>1\*</sup>, Wentao Hu<sup>1\*</sup>, Yanming Ma<sup>2</sup>, Yanbin Wang<sup>3</sup>, Zhisheng Zhao<sup>1</sup>, Bin Wen<sup>1</sup>, Julong He<sup>1</sup>, Zhongyuan Liu<sup>1</sup> & Yongjun Tian<sup>1</sup>

**Although diamond is the hardest material for cutting tools, poor thermal stability has limited its applications, especially at high temperatures. Simultaneous improvement of the hardness and thermal stability of diamond has long been desirable. According to the Hall–Petch effect<sup>1,2</sup>, the hardness of diamond can be enhanced by nanostructuring (by means of nanograined and nanotwinned microstructures), as shown in previous studies<sup>3–7</sup>. However, for well-sintered nanograined diamonds, the grain sizes are technically limited to 10–30 nm (ref. 3), with degraded thermal stability<sup>4</sup> compared with that of natural diamond. Recent success in synthesizing nanotwinned cubic boron nitride (nt-cBN) with a twin thickness down to ~3.8 nm makes it feasible to simultaneously achieve smaller nanosize, ultrahardness and superior thermal stability<sup>5</sup>. At present, nanotwinned diamond (nt-diamond) has not been fabricated successfully through direct conversions of various carbon precursors<sup>3,6,7</sup> (such as graphite, amorphous carbon, glassy carbon and C<sub>60</sub>). Here we report the direct synthesis of nt-diamond with an average twin thickness of ~5 nm, using a precursor of onion carbon nanoparticles at high pressure and high temperature, and the observation of a new monoclinic crystalline form of diamond coexisting with nt-diamond. The pure synthetic bulk nt-diamond material shows unprecedented hardness and thermal stability, with Vickers hardness up to ~200 GPa and an in-air oxidation temperature more than 200 °C higher than that of natural diamond. The creation of nanotwinned microstructures offers a general pathway for manufacturing new advanced carbon-based materials with exceptional thermal stability and mechanical properties.**

Diamond is the hardest, stiffest and least compressible crystalline material with exceptionally high thermal conductivity. Tools made of diamond are widely used for cutting and shaping hard substances such as stones, glasses and ceramics. However, diamond is energetically unstable relative to graphite under ambient conditions, with an inherent drawback of poor thermal stability. In air, the onset oxidation temperature is ~800 °C for natural diamond<sup>8,9</sup>, resulting in the severe wear of diamond tools at high temperatures.

The synthesis of materials harder than natural diamond has long been sought<sup>10</sup>. The Hall–Petch relation<sup>1,2</sup> offers a general pathway to enhancing hardness by decreasing characteristic size of microstructures (for example grain size or twin thickness). Nanograined diamond has been successfully synthesized through direct conversions of certain carbon precursors at high pressure and high temperature (HPHT)<sup>3,6,7</sup>. The pressure and temperature conditions<sup>6</sup> needed to synthesize nanograined diamonds are much higher than those for growing single-crystal diamonds in the industry. High pressure is necessary to control grain size effectively by suppressing atomic diffusion, which promotes growth. Nanograined diamonds synthesized from pure graphite at 2,300–2,500 °C and 12–25 GPa reach a grain size of 10–30 nm, with a high Knoop hardness of 110–140 GPa (ref. 3) but a reduced onset oxidation temperature of ~680 °C in air<sup>4</sup>. At lower temperatures (~1,800 °C), nanograined diamonds with a smaller grain size (5–10 nm) have been

synthesized from C<sub>60</sub>, amorphous carbon and glassy carbon, but Knoop hardness decreases significantly to 70–86 GPa (ref. 6). The observed hardness deficiency seems to originate from intergranular fracturing along poorly sintered grain boundaries, rather than the reverse Hall–Petch effect resulting from grain-boundary sliding<sup>6</sup>. Technically, the synthesis of well-sintered nanograined diamond while maintaining a smaller grain size remains a challenge.

Nanotwinning is an effective mechanism for acquiring a smaller characteristic size of microstructure, because twin boundaries possess lower excess energy than grain boundaries. It has been verified experimentally that, at nanoscale, twin boundaries show a hardening effect identical to those of grain boundaries for metals<sup>11,12</sup>. Ubiquitously nanotwinned structures have been introduced into superhard materials through the successful synthesis of nt-cBN with an average twin thickness of ~3.8 nm at HPHT<sup>5</sup>. These nt-cBN bulk samples have a superior combination of high hardness, high toughness and high thermal stability<sup>5</sup>. The synthesis of nt-diamond has not yet been reported but is highly desirable in view of the excellent performance of nt-cBN.

Experience in the synthesis of nt-cBN through an onion-like BN precursor suggested the use of onion carbon as precursor in the fabrication of nt-diamond. Onion carbon, a high-energy metastable carbon consisting of concentric graphite-like shells (Extended Data Fig. 1), is structurally similar to onion-like BN and can be produced in large amounts<sup>13</sup>. A high concentration of puckered layers and stacking faults in onion carbon may provide the key for the nucleation of nt-diamond at HPHT, as for nt-cBN<sup>5</sup>. In fact, isolated onion carbon particles have been observed to convert into diamond nanocrystals under intense electron irradiation even at ambient pressure<sup>14</sup>.

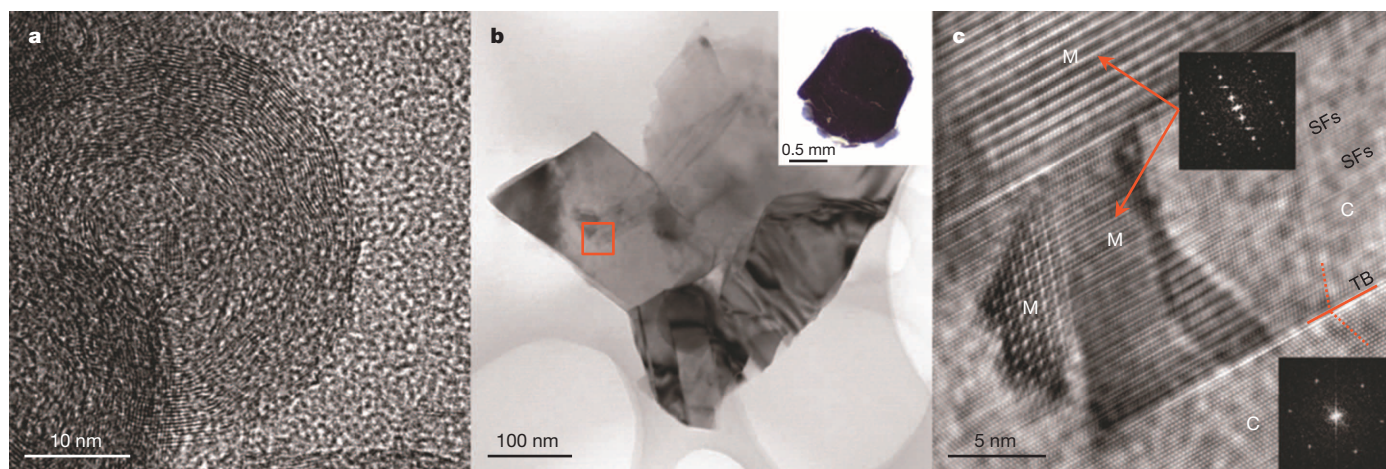
The onion carbon nanoparticles (~20–50 nm in diameter) used in our study were characterized by transmission electron microscopy (TEM) to contain numerous puckering and stacking faults (Fig. 1a). X-ray diffraction (XRD) characterization of the onion precursors and recovered samples after HPHT treatments is presented in Extended Data Figs 2 and 3. The inter-shell spacings of untreated onion carbon were centred on 0.3485 nm. When treated below 10 GPa and 2,000 °C, onion carbon retained the original nested crystal structure. Samples recovered from 10–15 GPa and 1,400–1,850 °C were black and opaque (Fig. 1b inset), and contained cubic diamond and an unidentified carbon phase. This latter phase has not been observed before and seems to be inherently related to the specific structural transformation of onion carbon precursors at HPHT. Transparent samples were recovered from 18–25 GPa and 1,850–2,000 °C, with pure cubic diamond as indicated by the XRD patterns. The synthetic temperature of cubic diamond from onion carbon was ~450 °C lower than that from graphite<sup>3,6</sup>, allowing easier industrial fabrication.

Typical TEM and high-resolution TEM (HRTEM) images of a black opaque sample (synthesized at 10 GPa and 1,850 °C) are shown in Fig. 1b, c and Extended Data Fig. 4a, b. Cubic diamond was the dominant phase, with lamellar {111} nanotwins. The new secondary carbon

<sup>1</sup>State Key Laboratory of Metastable Materials Science and Technology, Yanshan University, Qinhuangdao 066004, China. <sup>2</sup>State Key Laboratory for Superhard Materials, Jilin University, Changchun 130012, China. <sup>3</sup>Center for Advanced Radiation Sources, University of Chicago, Chicago, Illinois 60439, USA.

\*These authors contributed equally to this work.





**Figure 1 | Onion carbon nanoparticles and a bulk sample synthesized at 10 GPa and 1,850 °C.** **a**, HRTEM image of onion carbon nanoparticles. **b**, TEM image of the sample showing nanotwinned microstructure. Inset: photograph of the black opaque sample (~2 mm in diameter). **c**, HRTEM image of the area marked with the red box in **b**. Two adjacent cubic diamond

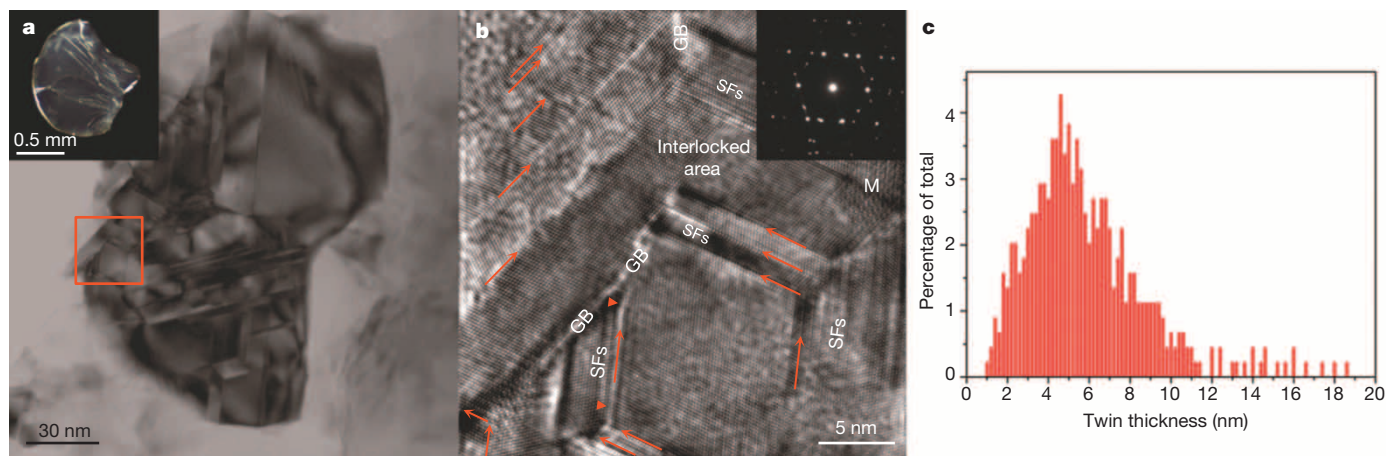
(C) domains form a {111} twin boundary (TB). Several M-diamond (M) domains are associated with cubic diamond twins containing stacking faults (SFs). Fast Fourier transforms of M-diamond and cubic diamond, shown in the upper and lower insets, respectively, indicate that lattices of M-diamond and cubic diamond are coherent.

phase was clearly seen with HRTEM. The  $d$  spacings deduced from selected-area electron diffraction (SAED) patterns (Extended Data Fig. 4d–f) and XRD data (Extended Data Table 1) of this new phase do not match any reported carbon phase. The new phase (denoted M-diamond) had a monoclinic structure with lattice parameters of  $a = 0.436$  nm,  $b = 0.251$  nm,  $c = 1.248$  nm and  $\beta = 90.9^\circ$ . All the C–C bonds were  $sp^3$  hybridized, as indicated by the electron energy loss spectrum measurements (Extended Data Fig. 4c), similar to those in cubic diamond. In the TEM images, thin, elongated (and occasionally polygonal) M-diamond domains intersected adjacent nanotwinned cubic diamond (C-diamond) domains, forming coherent boundaries parallel to the diamond (111) planes. The orientation relations between M-diamond (M) and C-diamond (C) as determined from SAED were  $M(001)//C(111)$  and  $M[010]/C[011]$  (Extended Data Fig. 4d–f).

The HRTEM images of a transparent pure nt-diamond sample (synthesized at 20 GPa and 2,000 °C; Fig. 2a inset) revealed that C-diamond contained a high density of lamellar {111} nanotwins (Fig. 2a, b). Unlike nt-cBN, in which individual nanograins can be clearly characterized<sup>5</sup>, high-angle grain boundaries in nt-diamond (Fig. 2b) were frequently

interrupted by interlocked areas where adjacent nanocrystals intersected and merged, making it difficult to determine individual nanograins unambiguously. The nanotwins were predominantly thinner than 10 nm. Figure 2c shows a twin thickness distribution derived from 444 nanotwins on the basis of HRTEM measurements. The average thickness, ~5 nm, is the smallest microstructural size so far achieved in diamonds. In our transparent nt-diamond samples, stacking faults were also observed in nanotwins (Fig. 2b and Extended Data Fig. 5). These stacking faults, due to extensive twinning, altered the stacking sequence of (111) planes in diamond<sup>15</sup> and produced weak shoulders of the strong (111) reflection in the XRD patterns (Extended Data Fig. 2). These observed planar faults together with the secondary phase of M-diamond also caused the asymmetries in both the (111) and the (220) peaks of diamond (Extended Data Figs 2 and 3).

A hardness value should be determined by the asymptotic region of the hardness–load curve<sup>10,16</sup>. We found that our samples reached asymptotic hardness at a load of 4.9 N (Fig. 3a). Vickers and Knoop hardnesses measured at 4.9 N for six different transparent pure nt-diamond samples (Fig. 3b and Extended Data Table 2) showed unprecedentedly high



**Figure 2 | A nt-diamond bulk sample synthesized at 20 GPa and 2,000 °C.** **a**, TEM image of nanotwinned microstructure. Inset: photograph of the transparent sample (~1 mm in diameter). **b**, HRTEM image of intersecting nanotwins (marked with the red box in **a**), viewed along the [101] zone axis of diamond. Lamellar {111} nanotwins, stacking faults and residual M-diamond (trace) are present. Twin boundaries are marked with red arrows. Grain

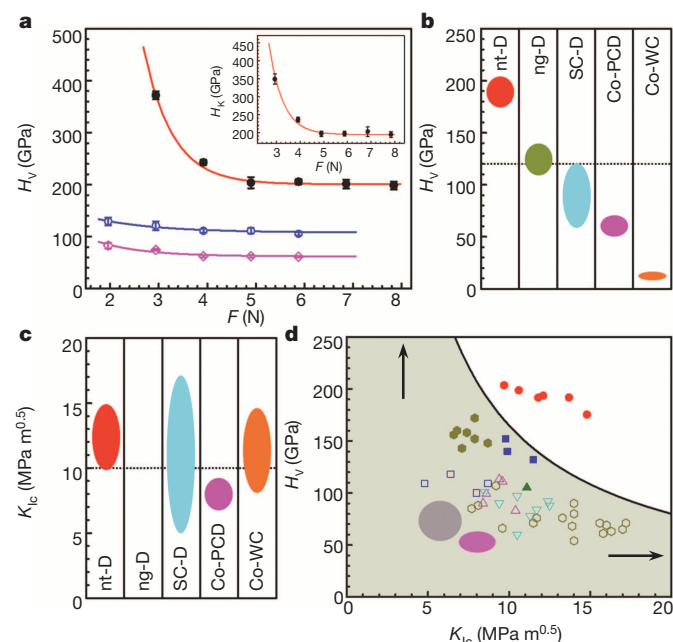
boundaries (GB) are interrupted by interlocked twins. Inset: SAED pattern corresponding to the central area of **a**. The four-fold-like pattern is from the twin domains with four different orientations. **c**, Thickness distribution of the nanotwins measured from HRTEM images. The average twin thickness is ~5 nm.

values: 175–203 and 168–196 GPa, respectively. Two high loads of 9.8 and 19.6 N were applied to create cracks for fracture toughness determination. The determined fracture toughness values ranged from 9.7 to 14.8 MPa m<sup>0.5</sup> (Fig. 3c and Extended Data Table 2). Meaningful indentation hardness can be measured reliably as long as the shear strength of the sample is smaller than the compressive strength of the diamond indenter<sup>16</sup>; this requirement was satisfied because no visible plastic deformation of indenter diamond tip was observed after measurements of hardness and fracture toughness (Extended Data Fig. 6). Both the achieved hardness and the trade-off between hardness and toughness of our nt-diamond samples are significantly superior to those of other popular tool materials, such as cobalt-bonded tungsten carbide (Co-WC)<sup>17</sup> and previously reported diamond-related materials<sup>18–21</sup> (Fig. 3b, d), yielding diamonds with unsurpassed mechanical properties. The simultaneous improvement in hardness and fracture toughness in our nt-diamond is intimately related to the ubiquitous nanotwinning microstructure. The presence of ultrafine nanotwins introduces extra hardening, which is probably due to both the Hall–Petch and quantum confinement effects at nanoscale<sup>5</sup>, while gliding of dislocations along densely distributed twin boundaries enhances fracture toughness<sup>22</sup>. Our results demonstrate that the old paradigm—the higher the hardness of a material, the lower

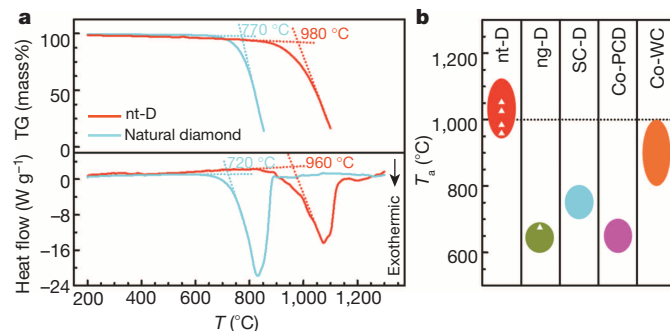
the fracture toughness—can be broken through processes of controlled nanotwinning in covalent materials.

The thermal stability of different pure nt-diamond samples was characterized by thermogravimetry curves measured in air. At a heating rate of 5 °C min<sup>−1</sup>, the onset oxidation temperatures of nt-diamond and natural diamond were ~980 and ~770 °C (Fig. 4a), respectively. Extended Data Fig. 7 compares the thermal stability of nt-diamond with other tool materials measured at a heating rate of 10 °C min<sup>−1</sup>. The onset oxidation temperature of nt-diamond (~1,056 °C) was again much higher than those of natural diamond (~805 °C), synthetic diamond powders (~725 °C), nanograined diamond (~680 °C)<sup>4</sup> and Co-WC (800–1,000 °C)<sup>17</sup> (Fig. 4b), and even rivalled that of ng-cBN (~1,187 °C)<sup>4</sup>. The oxidation of diamond generally has two simultaneous processes<sup>9</sup>, namely the oxidation of graphitized diamond and the oxidation of diamond itself. Previous experiments have shown that the oxidation temperature of graphite in air is ~50 °C lower than that of diamond<sup>23</sup>. According to the size-dependent pressure–temperature phase diagram derived from nanothermodynamic theory<sup>24</sup>, diamond becomes energetically stable over graphite at deep nanometre scale (~5 nm). This would certainly delay the graphitization of nt-diamond and would result in a higher oxidation temperature. Moreover, compressive stress introduces additional resistance to the oxidation of diamond. Given that the internal stress induced by nanotwinning boundaries increases with reduced twin thickness<sup>25</sup>, the oxidation process of nt-diamond may be retarded because of the presence of ultrafine nanotwins. Differential scanning calorimetry (DSC) measurements provided further evidence that thinner nanotwins result in an even higher oxidation temperature of ~1,300 °C (Extended Data Fig. 7a), consistent with the aforementioned speculation. Thus, both mechanical properties and thermal stability depend primarily on the achieved average twin thickness.

The successful syntheses of nt-diamond and nt-cBN show that nanotwinning microstructure is an effective route for simultaneously enhancing the hardness, fracture toughness and thermal stability of superhard materials. Our experimental results on nt-diamond further confirm that there is continuous hardening at nanotwinning sizes down to ~5 nm, which agrees with previous results on nt-cBN<sup>5</sup> but is in stark contrast with the sharp softening of metals at these nanometre scales. We therefore predict that pursuing microstructure with thinner nanotwin sizes may lead to findings of covalent materials with even superior properties. Here it may be instructive to estimate the lower limit of nanotwin thickness and the corresponding ultimately achievable hardness ( $H_{ua}$ ) of diamonds. If we take {111} twins in nt-diamond as the model system, the estimated minimal twin thickness,  $\lambda_{min}$ , is  $3d_{111} = 0.618$  nm



**Figure 3 | Typical mechanical properties of nt-diamond and its comparison with other tool materials.** **a**,  $H_V$  of nt-diamond and natural diamond crystal as a function of applied load ( $F$ ). Beyond 4.9 N,  $H_V$  decreases to the asymptotic values of ~200 GPa for nt-diamond (red line). For natural diamond crystals, our measured  $H_V$  values are ~110 GPa on the {110} face (blue line) and ~62 GPa on the {111} face (pink line). Error bars indicate 1 s.d. ( $n = 5$ ). Inset: plot of  $H_K$  against  $F$  for nt-diamond. **b**,  $H_V$  (**b**) and  $K_{IC}$  (**c**) for different tool materials, including nt-diamond (nt-D), nanograined diamond (ng-D; grain size 10–30 nm)<sup>3</sup>, single-crystal diamond (SC-D)<sup>18</sup>, cobalt-bonded polycrystalline diamond (Co-PCD)<sup>19</sup> and Co-WC<sup>17</sup>. **d**, Plot of  $H_V$  against  $K_{IC}$  for nt-diamond in comparison with available data on other forms of diamond. The data for nt-diamond are shown as solid red circles above the shaded envelope. The published data are from representative diamond materials, including type Ia natural SC-D (open upward triangles<sup>18</sup>), IIa natural SC-D (open squares<sup>18</sup>), HPHT-grown SC-D (open downward triangles<sup>18</sup>), CVD-grown SC-D (open hexagons<sup>18</sup>), annealed IIa natural SC-D (filled squares<sup>18</sup>), CVD-grown SC-D annealed at HPHT (filled hexagons<sup>18</sup>), Co-PCD (large grey circle<sup>19</sup>), CVD-grown PCD (large pink oval<sup>20</sup>) and aggregated diamond rod (Knoop hardness, filled upward triangle<sup>21</sup>). The hardness of nanograined diamond reaches 110–140 GPa (ref. 3), but no fracture toughness data were reported. Those data are therefore not included in the figure.



**Figure 4 | Typical thermal stability of a nt-diamond sample.** **a**, Comparison of the onset oxidation temperatures of a nt-diamond bulk sample (red) with a natural diamond crystal (cyan). Both thermogravimetry (TG; top) and DSC (bottom) curves were measured in air at a heating rate of 5 °C min<sup>−1</sup>. The onset oxidation temperature of the nt-diamond (980 °C from thermogravimetry or 960 °C from DSC) was more than 200 °C higher than that of the natural diamond (770 °C from thermogravimetry or 720 °C from DSC). **b**, Comparison of working temperatures ( $T_a$ ) in air of nt-diamond with other tool materials, including ng-D<sup>4</sup>, SC-D<sup>21</sup>, Co-PCD<sup>3</sup> and Co-WC<sup>17</sup>.



because the atomic stacking sequence along the  $\langle 111 \rangle$  direction of diamond (lattice parameter  $a = 0.3568$  nm) is ...ABCABC... (Extended Data Fig. 8). Assuming that the Hall–Petch effect is no longer applicable at such a scale<sup>26,27</sup>,  $H_{\text{ua}}$  for nt-diamond is estimated with the following formula according to our hardness model<sup>5,28</sup>:  $H_{\text{ua}} = H_0 + k_{\text{qc}}/\lambda_{\text{min}}$ , where  $H_0$  is the hardness of single-crystal diamond ( $\sim 90$  GPa) and  $k_{\text{qc}} = 211N_e^{1/3} = 187.7$  GPa nm is the quantum confinement hardening coefficient for a covalent crystal<sup>28</sup>, which is proportional to the valence electron density  $N_e$  (0.705, ref. 29). Thus,  $H_{\text{ua}}$  for nt-diamond is 394 GPa. This presents a technical challenge to synthesize nanotwinned microstructures with the required twin thickness to achieve such an exceptional hardness property.

Finally, the experimental HPHT conditions for synthesizing nt-diamond and nt-cBN are essentially identical. This opens up the possibility of manufacturing nt-diamond/nt-cBN composites. Such nanotwinned composites are expected to possess intermediate oxidation temperature and hardness between those of nt-diamond and nt-cBN but with greater fracture toughness as a result of the combined contributions from nanotwinning and composite effects.

## METHODS SUMMARY

We fabricated onion carbon particles with diameters of  $\sim 20$ – $50$  nm by using black carbon powders through an impinging-streams technology<sup>30</sup>. HPHT experiments were performed with a 10-MN double-stage large-volume multi-anvil system with the standard COMPRES 10/5 sample assembly consisting of a 10-mm spinel ( $\text{MgAl}_2\text{O}_4$ ) + MgO octahedron with a Re heater and a  $\text{LaCrO}_3$  thermal insulator. Temperature was measured with type C W–Re thermocouples, and pressure was estimated from previously obtained calibration curves at different temperatures for the multi-anvil apparatus<sup>5</sup>. Recovered samples were  $\sim 1$ – $2$  mm in diameter and 0.2–0.5 mm in height. Microstructures were investigated with a transmission electron microscope (JEM-2010) with an accelerating voltage of 200 kV. Component phases were identified by TEM and XRD (Cu  $K_\alpha$ ; D8 Discover). A microhardness tester (KB 5 BVZ) was used to measure  $H_V$  and  $K_{\text{IC}}$  with a diamond Vickers indenter as well as  $H_K$  with a diamond Knoop indenter.  $H_V$  was determined from  $H_V = 1,854.4F/L_1^2$ , where  $F$  (in newtons) is the applied load and  $L_1$  (in micrometres) is the arithmetic mean of the two diagonals of the Vickers indentation.  $H_K$  was determined from  $H_K = 14,228.9F/L_2^2$ , where  $L_2$  (in micrometres) is the longer diagonal of the Knoop indentation. Five hardness data points were obtained at each load, and the hardness values were determined from the asymptotic-hardness region.  $K_{\text{IC}}$  was calculated from  $K_{\text{IC}} = 0.016(E/H_V)^{0.5}F/C^{1.5}$  for radial cracks formed in the bulk nt-diamond sample<sup>18</sup>, where  $C$  (in micrometres) is the average length of the radial cracks measured from the indent centre, and  $E = 1,000$  GPa is Young's modulus of diamond<sup>18</sup>. The presented  $K_{\text{IC}}$  values were averaged over three data points determined at loads of 9.8 and 19.6 N. Oxidation resistance was studied by measuring thermogravimetry and DSC curves in air, using NETZSCH STA 449 C over the temperature range 20– $1,500^\circ\text{C}$ .

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 November 2013; accepted 15 April 2014.

- Hall, E. O. The deformation and ageing of mild steel. III. Discussion of results. *Proc. Phys. Soc. Lond. B* **64**, 747–753 (1951).
- Petch, N. J. The cleavage strength of polycrystals. *J. Iron Steel Inst.* **174**, 25–28 (1953).
- Irifune, T., Kurio, A., Sakamoto, S., Inoue, T. & Sumiya, H. Materials—ultrahard polycrystalline diamond from graphite. *Nature* **421**, 599–600 (2003).
- Solozhenko, V. L., Kurakevich, O. O. & Le Godec, Y. Creation of nanostructures by extreme conditions: high-pressure synthesis of ultrahard nanocrystalline cubic boron nitride. *Adv. Mater.* **24**, 1540–1544 (2012).

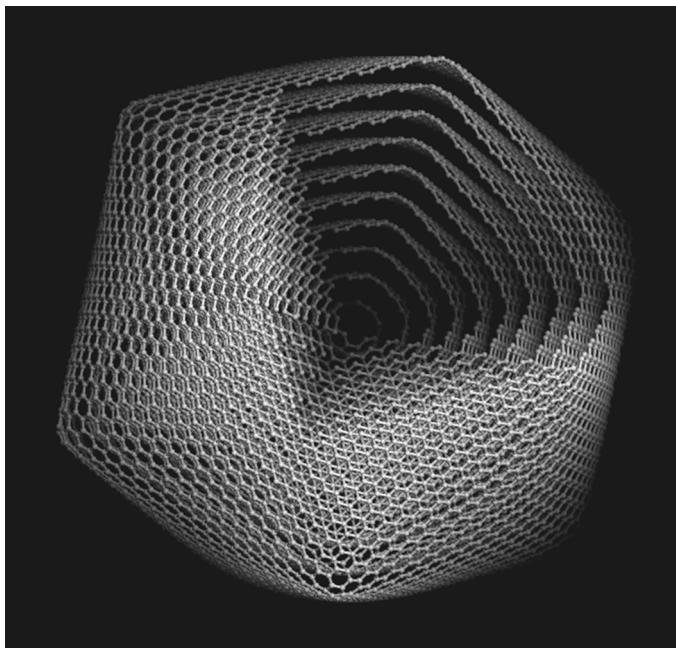
- Tian, Y. *et al.* Ultrahard nanotwinned cubic boron nitride. *Nature* **493**, 385–388 (2013).
- Sumiya, H. & Irifune, T. Hardness and deformation microstructures of nano-polycrystalline diamonds synthesized from various carbons under high pressure and high temperature. *J. Mater. Res.* **22**, 2345–2351 (2007).
- Sumiya, H. & Harano, K. Distinctive mechanical properties of nano-polycrystalline diamond synthesized by direct conversion sintering under HPHT. *Diamond Relat. Mater.* **24**, 44–48 (2012).
- Sun, Q. & Alam, M. Relative oxidation behavior of chemical vapor deposited and type IIa natural diamonds. *J. Electrochem. Soc.* **139**, 933–936 (1992).
- Johnson, C. E., Bennett, J. M. & Nadler, M. P. Oxidation of diamond windows. *J. Mater. Res.* **10**, 2555–2563 (1995).
- Chaudhri, M. M. & Lim, Y. Y. Harder than diamond? Just fiction. *Nature Mater.* **4**, 4 (2005).
- Lu, L., Shen, Y. F., Chen, X. H. & Lu, K. Ultrahigh strength and high electrical conductivity in copper. *Science* **304**, 422–426 (2004).
- Lu, L., Chen, X., Huang, X. & Lu, K. Revealing the maximum strength in nanotwinned copper. *Science* **323**, 607–610 (2009).
- Choucair, M. & Stride, J. A. The gram-scale synthesis of carbon onions. *Carbon* **50**, 1109–1115 (2012).
- Banhart, F. & Ajayan, P. M. Carbon onions as nanoscopic pressure cells for diamond formation. *Nature* **382**, 433–435 (1996).
- Silva, F., Bénédict, F., Bruno, P. & Gicquel, A. Formation of  $\langle 110 \rangle$  texture during nanocrystalline diamond growth: an X-ray diffraction study. *Diamond Relat. Mater.* **14**, 398–403 (2005).
- Tian, Y. *et al.* Controversy about ultrahard nanotwinned cBN reply. *Nature* **502**, E2–E3 (2013).
- Upadhyaya, G. S. Materials science of cemented carbides—an overview. *Mater. Des.* **22**, 483–489 (2001).
- Yan, C. S. *et al.* Ultrahard diamond single crystals from chemical vapor deposition. *Phys. Status Solidi A* **201**, R25–R27 (2004).
- Lammer, A. Mechanical properties of polycrystalline diamonds. *Mater. Sci. Technol.* **4**, 949–955 (1988).
- Sussmann, R. S. *et al.* CVD diamond windows for infrared synchrotron applications. *Nuovo Cimento D* **20**, 503–525 (1998).
- Dubrovinskaia, N., Dub, S. & Dubrovinsky, L. Superior wear resistance of aggregated diamond nanorods. *Nano Lett.* **6**, 824–826 (2006).
- Lu, K., Lu, L. & Suresh, S. Strengthening materials by engineering coherent internal boundaries at the nanoscale. *Science* **324**, 349–352 (2009).
- Joshi, A., Nimmagadda, R. & Herrington, J. Oxidation kinetics of diamond, graphite, and chemical vapor deposited diamond films by thermal gravimetry. *J. Vac. Sci. Technol. A* **8**, 2137–2142 (1990).
- Yang, C. C. & Li, S. Size-dependent temperature–pressure phase diagram of carbon. *J. Phys. Chem. C* **112**, 1423–1426 (2008).
- Weissmüller, J. & Cahn, J. W. Mean stresses in microstructures due to interface stresses: a generalization of a capillary equation for solids. *Acta Mater.* **45**, 1899–1906 (1997).
- Yip, S. Nanocrystals: the strongest size. *Nature* **391**, 532–533 (1998).
- Li, X., Wei, Y., Lu, L., Lu, K. & Gao, H. Dislocation nucleation governed softening and maximum strength in nano-twinned metals. *Nature* **464**, 877–880 (2010).
- Tian, Y., Xu, B. & Zhao, Z. Microscopic theory of hardness and design of novel superhard crystals. *Int. J. Refract. Met. Hard Mater.* **33**, 93–106 (2012).
- Gao, F. M. *et al.* Hardness of covalent crystals. *Phys. Rev. Lett.* **91**, 015502 (2003).
- Tamir, A. *Impinging-stream Reactors: Fundamentals and Applications* (Elsevier, 1994).

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (51121061), the Ministry of Science and Technology of China (2011CB808205 and 2010CB731605), the National Natural Science Foundation of China (51332005, 51172197, 11025418 and 91022029) and the US National Science Foundation (EAR-0968456).

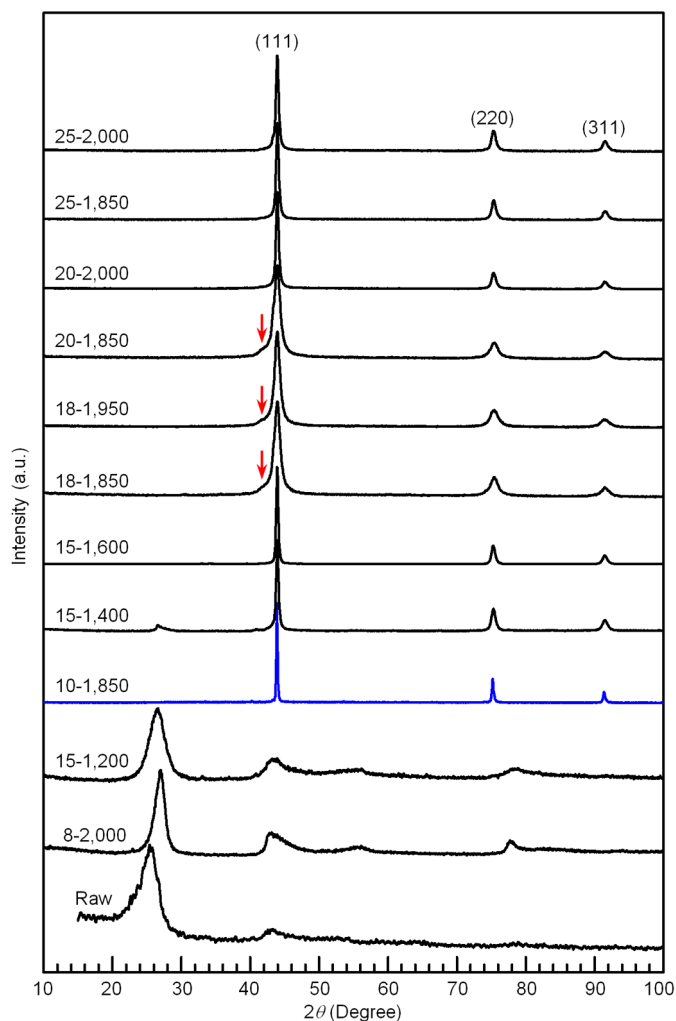
**Author Contributions** Y.J.T. conceived the project. Y.J.T., D.L.Y., B.X. and Y.B.W. designed the experiments. Q.H. synthesized onion carbon precursors. Q.H., D.L.Y., B.X., Y.J.T., Y.B.W. and Z.S.Z. performed the HPHT experiments, W.T.H. performed TEM observations, and B.W. performed molecular dynamics simulations. Y.J.T., B.X., D.L.Y., Y.M.M., Y.B.W., J.L.H. and Z.Y.L. analysed the data. Y.J.T., B.X., Y.M.M. and Y.B.W. co-wrote the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.J.T. (fhcl@ysu.edu.cn).



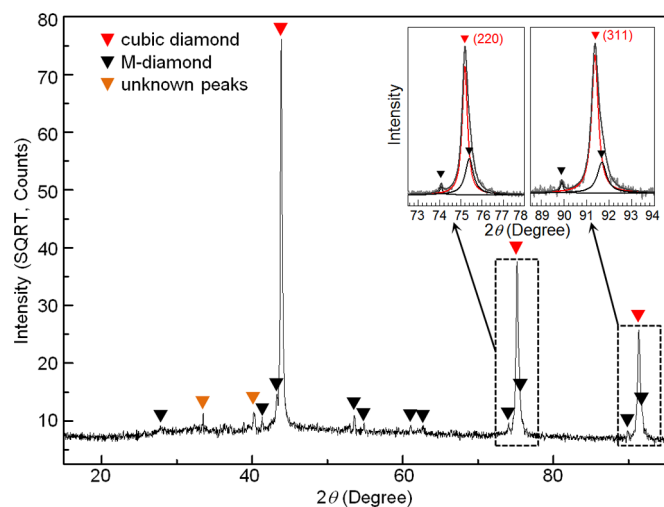


**Extended Data Figure 1 | Schematic icosahedral model of a ten-shell onion carbon.** The icosahedral-quasicrystal-like model of an onion carbon particle was relaxed from a nested buckyunion of  $C_{60}$ ,  $C_{240}$ ,  $C_{540}$ ,  $C_{960}$ ,  $C_{1,500}$ ,  $C_{2,160}$ ,  $C_{2,940}$ ,  $C_{3,840}$ ,  $C_{4,860}$  and  $C_{6,000}$ . This model was constructed with the same classical molecular dynamics technique as that used in our previous work<sup>5</sup>. The spacings between adjacent shells in the model vary from  $\sim 0.300$  nm to  $\sim 0.340$  nm.



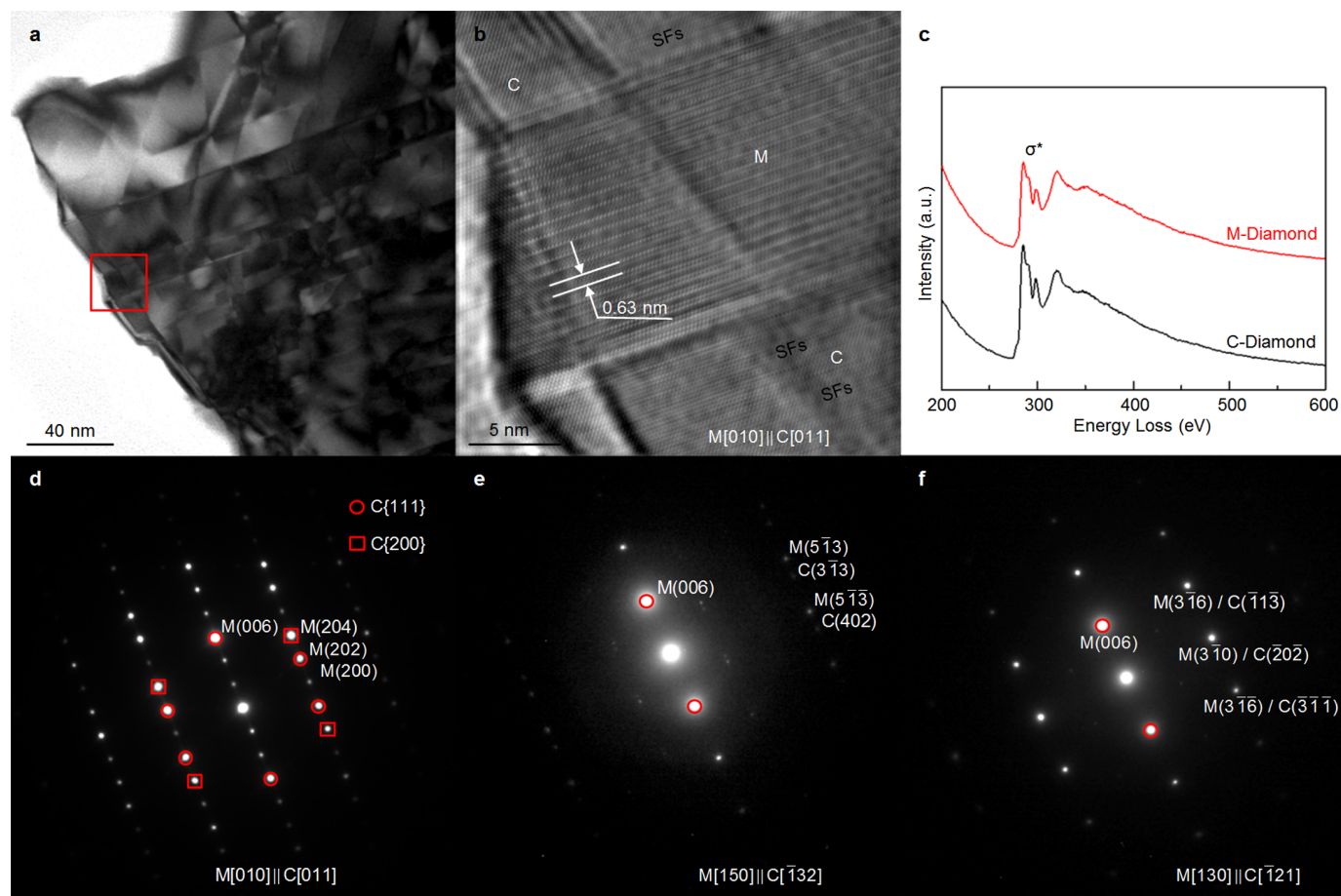
#### Extended Data Figure 2 | Phase transformation of onion carbon compacts at HPHT.

XRD patterns of onion carbon precursor (Raw) and seven samples recovered from different conditions indicated by  $P$  (in GPa)– $T$  (in  $^\circ\text{C}$ ) pairs. The inter-shell spacing of the starting onion carbon nanoparticles is  $\sim 0.3485$  nm. For the two samples recovered from 8 GPa/2,000  $^\circ\text{C}$  and 15 GPa/1,200  $^\circ\text{C}$ , the onion carbon structure does not show significant alteration except that the inter-shell spacing decreases to 0.3305 and 0.3361 nm, respectively. Cubic diamond appears when the applied pressure is more than 10 GPa and temperature is more than 1,400  $^\circ\text{C}$ , with an accompanying new carbon phase recognized in the black opaque samples synthesized at 1,850  $^\circ\text{C}$  or below. A small amount of residual onion carbon can be detected in the sample recovered from 15 GPa/1,400  $^\circ\text{C}$ . At pressures of 18–25 GPa and temperatures of 1,850–2,000  $^\circ\text{C}$ , the recovered samples changed from translucent to transparent, and only the diffraction peaks of cubic diamond can be seen in XRD patterns. Weak shoulders of the (111) peaks of diamond (red arrows) appear in three samples synthesized at pressures of 18–20 GPa and temperatures of 1,850–1,950  $^\circ\text{C}$ . Asymmetry in the (111) and (220) peaks of diamond was often observed in the samples synthesized at pressures below 20 GPa and temperatures below 1,950  $^\circ\text{C}$ .



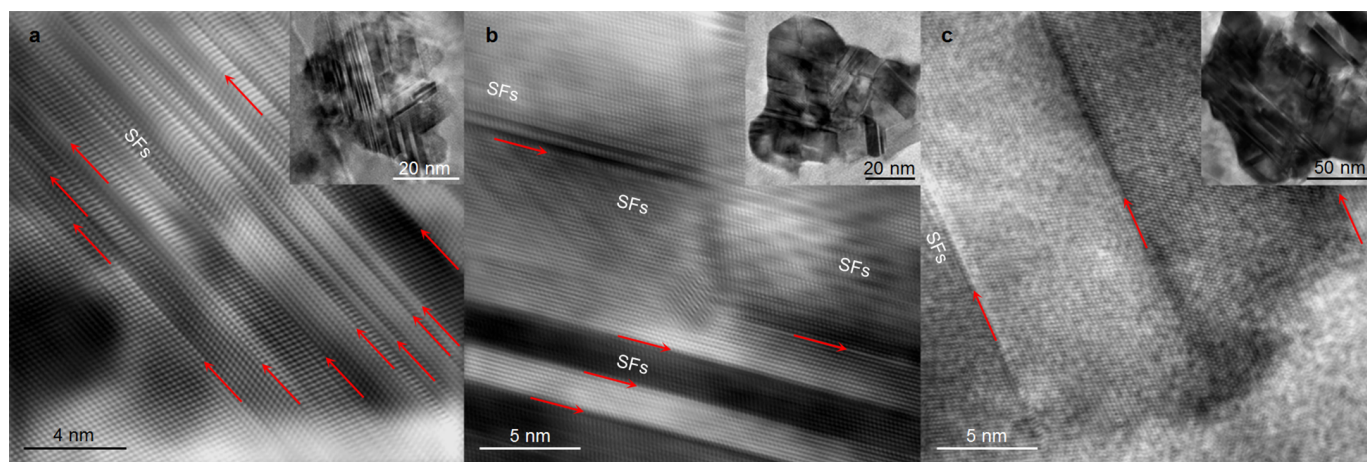
**Extended Data Figure 3 | XRD patterns of a sample recovered from 10 GPa and 1,850 °C.** All the recorded  $d$  spacings of visible diffraction peaks are listed in Extended Data Table 1. Insets: two peaks overlapping the cubic diamond reflections. Most of these extra reflections can be indexed with a monoclinic structure (M-diamond) as shown in Extended Data Table 1.





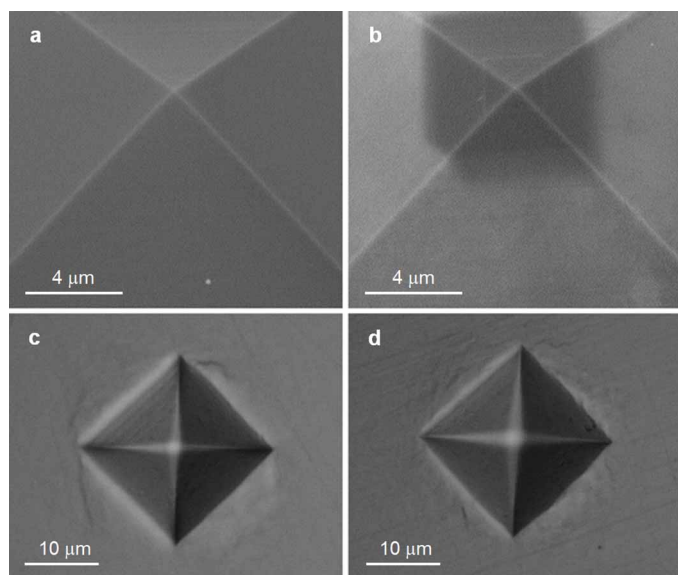
**Extended Data Figure 4 | TEM images, electron energy loss spectrum (EELS) and SAED measurements on a sample recovered from 10 GPa and 1,850 °C.** **a**, TEM image showing interlaced twins. **b**, HRTEM image corresponding to the area in the red box in **a**. A monoclinic M-diamond (M) domain is observed between two cubic diamond (C) domains. **c**, EELS spectra of M and C phases. All the C–C bonds are  $sp^3$  hybridized in both M and C

phases. **d–f**, SAED patterns along the  $[010]$ ,  $[150]$  and  $[130]$  zone axes of M, respectively, recorded by rotating an M crystal.  $(111)$  and  $(200)$  spots of the twinned C phase, overlapping with some spots of the M phase as a result of coherent growth, are marked by red circles and boxes, respectively. The determined orientation relations between M and C phases are  $M(001) \parallel C(111)$  and  $M[010] \parallel C[011]$ .



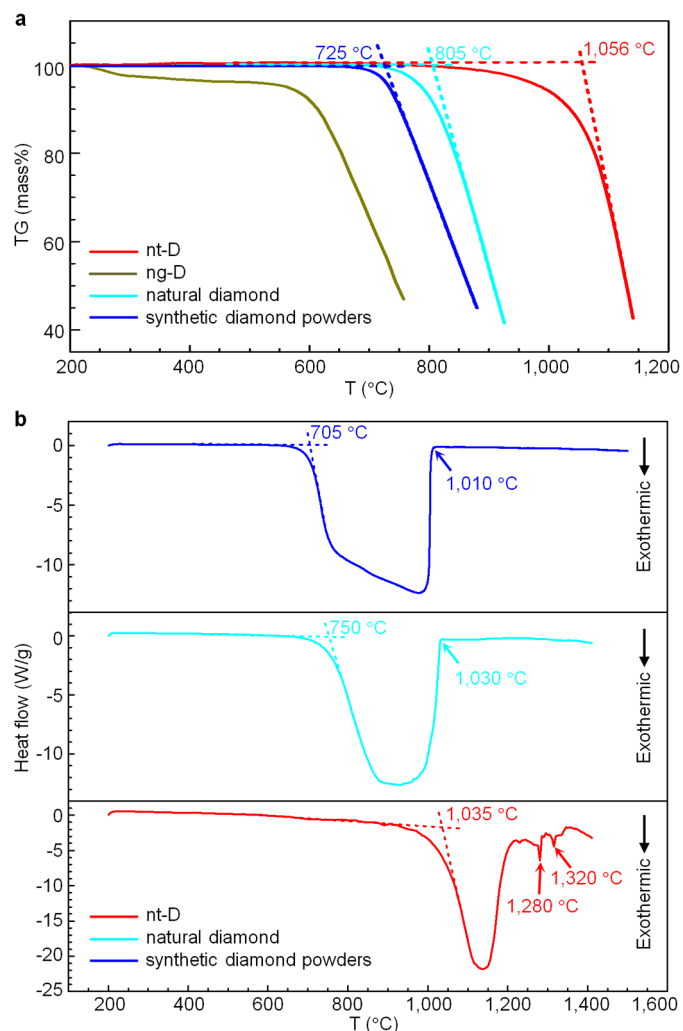
**Extended Data Figure 5 | HRTEM observations of three nt-diamond bulk samples synthesized at different HPHT conditions.** a–c, HRTEM and corresponding TEM (inset) images of three representative samples, O-366 (a), P-368 (b) and M-363 (c) as listed in Extended Data Table 2. TBs are marked with red arrows. The measured average twin thicknesses are  $\sim 5.2$  nm for sample P-368,  $\sim 5.4$  nm for sample O-366 and  $\sim 7.9$  nm for sample M-363; the smaller the average twin thickness, the higher the hardness. The full width at half-maximum (FWHM) of the (111) peak is mainly related to the nanograin size: samples O-366 and P-368 have a larger FWHM as a result of their smaller nanograin size. Both pressure and temperature can promote the phase

transformation of onion carbon to diamond. The probability of stacking faults and the volume fraction of M-diamond decrease with elevated synthesis temperature and pressure, as confirmed by our HRTEM observation. The abundant stacking faults in the nanotwins result in the appearance of a shoulder near the (111) peak (Extended Data Fig. 2), for example in the XRD pattern of sample O-366. The asymmetries of the (111) and (220) peaks of diamond shown in Extended Data Fig. 2 can be attributed to planar faults and the secondary phase in microstructure. On the one hand, a twin fault can itself produce peak asymmetry; on the other, M-diamond also contributes to peak asymmetry because of peak overlap, as demonstrated in Extended Data Fig. 3.

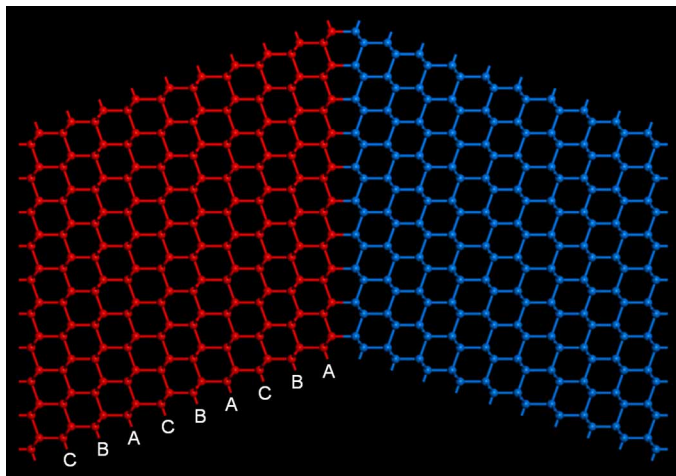


**Extended Data Figure 6 | Comparison of Vickers indenter tip before and after hardness and fracture toughness tests of nt-diamond.** **a, b,** Scanning electron microscopy images of the square pyramid diamond tip before (**a**) and after (**b**) the tests of nt-diamond. A load of 9.8 N was used during the hardness and toughness tests. As shown in **b**, the indenter, with a dark imprint of  $\sim 6.9 \mu\text{m} \times \sim 6.9 \mu\text{m}$  on the tip matching the permanent indentation on the tested nt-diamond, shows no visible plastic deformation. **c, d,** Photographs of indentations on the standard calibration block equipped by microhardness tester KB 5 BVZ. The indentations were formed at a load of 1.96 N before (**c**) and after (**d**) the tests, with the same tip as shown in **a** and **b**. The indenter tip produced an almost identical indentation (or standard hardness value) on the calibration block after the nt-diamond tests. These calibration results ensured the accuracy, repeatability and reliability of the unprecedented hardness and exceptional toughness values of nt-diamond reported in the present study.





**Extended Data Figure 7 | Comparison of in-air oxidation resistance of bulk nt-diamond with other diamonds measured at a heating rate of  $10\text{ }^{\circ}\text{C min}^{-1}$ .** **a**, Comparison of the onset oxidation temperatures determined from measured thermogravimetry curves. The onset temperature was  $\sim 1,056\text{ }^{\circ}\text{C}$  for a bulk nt-diamond,  $\sim 805\text{ }^{\circ}\text{C}$  for a natural diamond crystal,  $\sim 725\text{ }^{\circ}\text{C}$  for synthetic diamond powders and  $\sim 680\text{ }^{\circ}\text{C}$  for a nanograined diamond<sup>4</sup>. **b**, Comparison of the onset oxidation temperatures determined from the exothermic trough in the measured heat flow curves of DSC. The onset temperature was  $\sim 1,035\text{ }^{\circ}\text{C}$  for the nt-diamond,  $\sim 750\text{ }^{\circ}\text{C}$  for the natural diamond and  $\sim 705\text{ }^{\circ}\text{C}$  for the synthetic diamond. The exothermic peaks located at  $1,280\text{ }^{\circ}\text{C}$  and  $1,320\text{ }^{\circ}\text{C}$  for the nt-diamond were probably due to the presence of finer nanotwins. The above-measured oxidation temperatures are consistent with those determined from the corresponding thermogravimetry curves.



**Extended Data Figure 8 | Atomic arrangements of a  $\{111\} \Sigma = 3$  twin boundary in cubic diamond.** The twin boundary is projected along the  $\langle 011 \rangle$  direction. Because of the stacking sequence of ABC for diamond structure, the minimum twin thickness is  $3d_{111}$ , where  $d_{111}$  is the planar distance along the direction of  $\langle 111 \rangle$  in the unit cell of cubic diamond.

**Extended Data Table 1 | Comparison of  $d$  spacings ( $d_{\text{obs}}$ ) observed from XRD and SAED with those of proposed M-diamond structure and cubic diamond**

Sample			M-diamond		Cubic diamond JCPDS6-675		
$d_{\text{obs}}$ (nm) (XRD)	$hkl$	$d_{\text{obs}}$ (nm) (SAED)	$d_{\text{cal}}$ (nm)	$hkl$	$d$ (nm)	$hkl$	$hkl$
		0.625	0.6239	0 0 2			
0.3217	0.25	0.313	0.3120	0 0 4			
0.2672*	1.05						
0.2239*	0.78						
0.2182	0.50	0.218	0.2179	2 0 0			
0.2176	0.40	0.212	0.2142	1 -1 1			
0.2087	1.56	0.209	0.2080	0 0 6			
0.2062	100	0.206	0.2068	2 0 -2	0.206	1 1 1	100
		0.205	0.2048	2 0 2			
		0.180	0.1800	2 0 -4			
0.1707	0.92		0.1774	2 0 4			
0.1671	0.52		0.1647	1 1 -5			
0.1515	0.27	0.154	0.1517	2 0 -6			
0.1481	0.28		0.1493	2 0 6			
0.1279	0.53	0.128	0.1278	2 0 -8			
0.1263	23.34				0.1261	2 2 0	25
0.1260	2.58	0.126	0.1259	2 0 8			
		0.126	0.1258	3 -1 0			
		0.110	0.1090	4 0 0			
0.1090	0.25	0.109	0.1090	2 0 -10			
		0.108	0.1083	3 -1 -6			
0.1077	9.36	0.108	0.1076	2 0 10			
0.1074	2.49	0.108	0.1076	4 0 -2	0.10754	3 1 1	16
		0.107	0.1071	4 0 2			
		0.107	0.1070	3 -1 6			
		0.103	0.1034	4 0 -4			
		0.102	0.1024	4 0 4			
		0.097	0.0972	4 0 -6			
		0.095	0.0961	4 0 6			
		0.091	0.0900	4 0 -8			
		0.089	0.0887	4 0 8	0.08916	4 0 0	8
		0.082	0.0818	5 -1 -3	0.08182	3 3 1	16
		0.081	0.0807	3 -1 12			
		0.081	0.0806	5 -1 3			
		0.080	0.0796	3 -1 -12			
		0.075	0.0744	5 -1 -7			
					0.07281	4 2 2	
					0.06864	5 1 1	

The sample was synthesized at 10 GPa and 1,850 °C. The  $d_{\text{cal}}$  values were calculated with the monoclinic structural parameters  $a = 0.436$  nm,  $b = 0.251$  nm,  $c = 1.248$  nm and  $\beta = 90.9^\circ$ . Asterisks indicate unknown peaks.



**Extended Data Table 2 | Vickers hardness  $H_V$  (GPa), Knoop hardness  $H_K$  (GPa) and fracture toughness  $K_{Ic}$  (MPa m<sup>0.5</sup>) for six transparent pure (XRD standard) nt-diamond bulk samples**

Samples	P(GPa) /T (°C)	FWHM (°)	$H_V \pm \text{s.d.}$ (n=5)	$H_K \pm \text{s.d.}$ (n=5)	$K_{Ic} \pm \text{s.d.}$ (n=3)
W-377	25/2,000	0.45	193.6 $\pm$ 21.7	182.1 $\pm$ 6.4	12.1 $\pm$ 0.5
M-363	25/2,000	0.43	175.4 $\pm$ 18.2	168.2 $\pm$ 6.8	14.8 $\pm$ 3.6
O-366	18/1,850	0.98	191.8 $\pm$ 9.3	190.1 $\pm$ 14.7	13.7 $\pm$ 4.2
V-376	20/2,000	0.46	203.6 $\pm$ 12.0	196.6 $\pm$ 8.2	9.7 $\pm$ 2.6
P-368	20/1,850	1.04	198.7 $\pm$ 32.5	190.9 $\pm$ 3.2	10.6 $\pm$ 1.1
S-379	20/1,850	0.43	191.7 $\pm$ 25.6	183.6 $\pm$ 6.8	11.8 $\pm$ 0.5

$H_V$  and  $H_K$  values were measured at a fixed load of 4.9 N. The  $K_{Ic}$  values were measured at loads of 9.8 and 19.6 N. Error bars indicate 1 s.d. ( $n = 5$  for  $H_V$  and  $H_K$ , and  $n = 3$  for  $K_{Ic}$ ). The FWHMs of (111) peaks in the XRD patterns of nt-diamond samples are also listed.

# Increased frequency of extreme Indian Ocean Dipole events due to greenhouse warming

Wenju Cai<sup>1,2</sup>, Agus Santoso<sup>3</sup>, Guojian Wang<sup>2,1</sup>, Evan Weller<sup>1</sup>, Lixin Wu<sup>2</sup>, Karumuri Ashok<sup>4</sup>, Yukio Masumoto<sup>5,6</sup> & Toshio Yamagata<sup>7</sup>

The Indian Ocean dipole is a prominent mode of coupled ocean–atmosphere variability<sup>1–4</sup>, affecting the lives of millions of people in Indian Ocean rim countries<sup>5–15</sup>. In its positive phase, sea surface temperatures are lower than normal off the Sumatra–Java coast, but higher in the western tropical Indian Ocean. During the extreme positive-IOD (pIOD) events of 1961, 1994 and 1997, the eastern cooling strengthened and extended westward along the equatorial Indian Ocean through strong reversal of both the mean westerly winds and the associated eastward-flowing upper ocean currents<sup>1,2</sup>. This created anomalously dry conditions from the eastern to the central Indian Ocean along the Equator and atmospheric convergence farther west, leading to catastrophic floods in eastern tropical African countries<sup>13,14</sup> but devastating droughts in eastern Indian Ocean rim countries<sup>8–10,16,17</sup>. Despite these serious consequences, the response of pIOD events to greenhouse warming is unknown. Here, using an ensemble of climate models forced by a scenario of high greenhouse gas emissions (Representative Concentration Pathway 8.5), we project that the frequency of extreme pIOD events will increase by almost a factor of three, from one event every 17.3 years over the twentieth century to one event every 6.3 years over the twenty-first century. We find that a mean state change—with weakening of both equatorial westerly winds and eastward oceanic currents in association with a faster warming in the western than the eastern equatorial Indian Ocean—facilitates more frequent occurrences of wind and oceanic current reversal. This leads to more frequent extreme pIOD events, suggesting an increasing frequency of extreme climate and weather events in regions affected by the pIOD.

In austral winter and spring, southeasterly trade winds that feed the tropical convergence zone near the maritime continent are a feature of the southern tropical Indian Ocean. During a pIOD event, an initial cooling off Sumatra–Java, the eastern pole of the Indian Ocean dipole, suppresses local convection, inducing easterly wind anomalies and a shallowing thermocline. This promotes upwelling that in turn reinforces the initial cooling<sup>1,2,18</sup>, a process referred to as Bjerknes feedback. The growth of cool anomalies causes a northward extension of the southeasterly trade winds<sup>1,2,16</sup>, with anomalous easterlies along the equatorial Indian Ocean (Fig. 1a), where weak westerlies normally prevail. The change in wind promotes convergence, rainfall and warm anomalies in the equatorial western Indian Ocean. The altered circulations induce droughts and bushfires in eastern Asia and Australia<sup>5–8</sup>, floods in parts of the Indian subcontinent<sup>11</sup> and eastern Africa<sup>13,14</sup>, coral reef death across western Sumatra<sup>12</sup>, and malaria outbreaks in eastern Africa<sup>15</sup>. During extreme pIOD events, as occurred in 1961, 1994 and 1997, the anomalies, particularly the anomalous equatorial easterlies, are far stronger (Fig. 1b), with commensurately greater impacts. During the 1997 event, devastating floods in Somalia, Ethiopia, Kenya, Sudan and Uganda caused several thousand deaths and displaced hundreds of thousands of people. In contrast, Indonesia suffered severe droughts and wildfires<sup>2,16,17</sup> made worse by the developing 1997 El Niño; the associated smoke

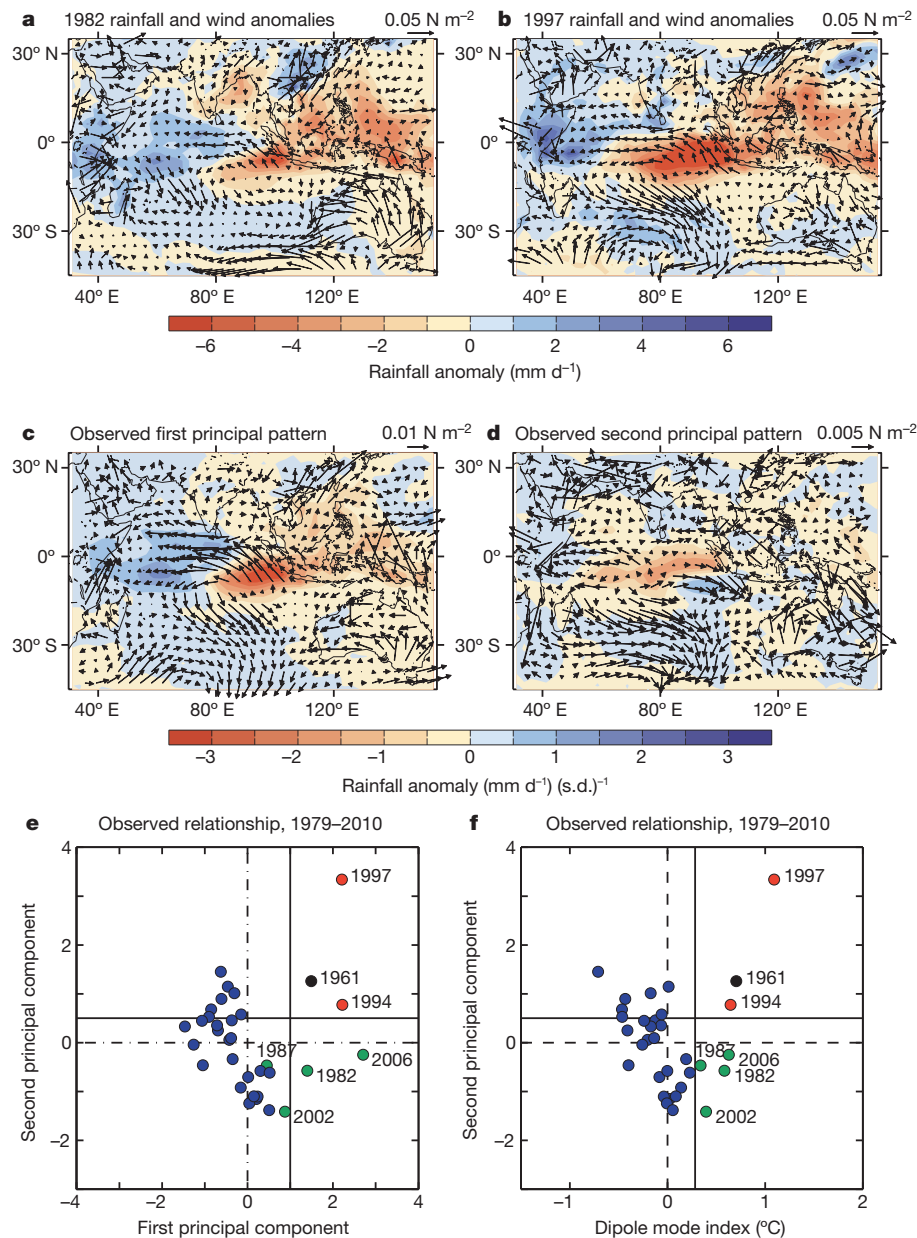
and haze caused severe health problems to tens of millions of people in Indonesia and surrounding countries<sup>9,10</sup>.

These dramatic impacts call for an urgent investigation into whether extreme pIOD events will change in a warmer climate. Recent studies have shown that greenhouse warming leads to a mean state change in the equatorial Indian Ocean with an easterly wind trend and a faster warming rate in the west than in the east, but referenced to the evolving mean state there is no detectable change in either the overall frequency or amplitude of pIOD events<sup>19–21</sup>. Here, using a suite of distinct process-based indicators, we show that there is in fact a significant increase in the frequency of the extreme pIOD events under greenhouse warming.

We characterize the observed extreme pIOD events in terms of their contrast with moderate events, focusing on austral spring, the season in which the IOD usually peaks. During extreme pIOD events, the cooling off Sumatra is intensified by the large equatorial easterly anomalies through generation of equatorial and coastally trapped upwelling Kelvin waves<sup>22,23</sup>, enhanced evaporation<sup>3</sup>, and a weakening of the mean eastward oceanic flows that transport heat eastward towards Sumatra<sup>24</sup>. The anomalous convergence in the west, marked by increased rainfall and temperature, is amplified through a series of processes: reduced wind speed and evaporation associated with the downstream extension of the southeasterly trades; a deeper thermocline caused by the weaker eastward ocean heat transport along the Equator<sup>24</sup>; and generation of equatorial downwelling Rossby waves<sup>3,4</sup>. The warming in the west and cooling in the east in turn strengthens the equatorial easterly anomalies, introducing a positive feedback along the Equator that operates in addition to the Bjerknes feedback centred off Sumatra–Java. The equatorial positive feedback, which is far stronger during extreme pIOD events, leads to stronger equatorial cooling (Extended Data Fig. 1a), and reversal of the equatorial winds and ocean currents so that they flow towards the west (Extended Data Fig. 2f). This creates a zone of atmospheric subsidence along the Equator characterized by low rainfall and colder sea surface temperatures (SSTs) that extend much farther to the west than during moderate pIOD events (Fig. 1; Extended Data Fig. 1a).

A heat budget analysis for the eastern-to-central equatorial Indian Ocean during the IOD developing phase (July to October; Extended Data Figs 1 and 2 and Methods) clearly indicates the 1961, 1994 and 1997 events to be the most extreme pIODs. The growth of equatorial SST anomalies during these three events is dominated by nonlinear processes involving zonal current anomalies. In particular, the nonlinear zonal advection, that is, the product of the anomalous west-minus-east SST gradient with the anomalous zonal currents (dark red bar, Extended Data Fig. 1c), sets these three events apart from the rest. Essentially, the equatorial positive feedback enhances anomalies of westward-flowing equatorial winds and currents, allowing for an eventual reversal. This nonlinear process can be parameterized by the product of the equatorial easterly anomalies, which drive the current, and the dipole mode index (DMI)<sup>1</sup>, which measures the west-minus-east SST gradient (see Methods). Such nonlinearity also occurs

<sup>1</sup>CSIRO Marine and Atmospheric Research, Aspendale, Victoria, 3195 Australia. <sup>2</sup>Physical Oceanography Laboratory, Qingdao Collaborative Innovation Center of Marine Science and Technology, Ocean University of China, Qingdao, 266003 China. <sup>3</sup>Climate Change Research Centre and ARC Centre of Excellence for Climate System Science, The University of New South Wales, Sydney, 2052 Australia. <sup>4</sup>Centre for Climate Change Research, Indian Institute of Tropical Meteorology Pashan, Pune 411 008, India. <sup>5</sup>Department of Earth and Planetary Science, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. <sup>6</sup>Climate Variation Predictability and Applicability Research Program, Japan Agency for Marine–Earth Science and Technology (JAMSTEC), 3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan. <sup>7</sup>Application Laboratory, JAMSTEC, 3173-25 Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan.



**Figure 1 | Comparison of moderate and extreme pIOD and identification of extreme pIOD events.** **a, b,** September–November average rainfall (shading, in units of mm day<sup>-1</sup>) and wind stress (vector scale is shown in the top right corner for each panel) anomalies associated with a moderate (1982) and extreme (1997) pIOD. **c, d,** Principal variability patterns of rainfall obtained by applying a statistical and signal processing method, EOF analysis, to a satellite-era rainfall anomaly data from the Global Precipitation Climatology Project version 2 (see Methods), in the equatorial region (10°S–10°N, 40°E–100°E). The associated rainfall and wind stress vectors from reanalysis data (see Methods) are presented as linear regression onto the EOF time series. The colour scale indicates rainfall in mm day<sup>-1</sup> per 1 s.d. change; blue or red

contours indicate increased or decreased rain. Note the different vector scales in **c** and **d**. **e,** Relationship between the two principal component time series. Values for 1961 are obtained by regressing the rainfall anomaly pattern from a reanalysis onto the EOF1 and EOF2 pattern (see Methods). An extreme pIOD event (red dots) is defined as when the first principal component is greater than 1 s.d. and the second principal component is greater than 0.5 s.d. A moderate pIOD event (green dots) is determined from a detrended DMI<sup>1</sup> when its amplitude is greater than 0.75 s.d. other than the 1994 and 1997 events. Negative IOD and neutral years are indicated with blue dots. **f,** Relationship between the second principal component time series and the DMI.

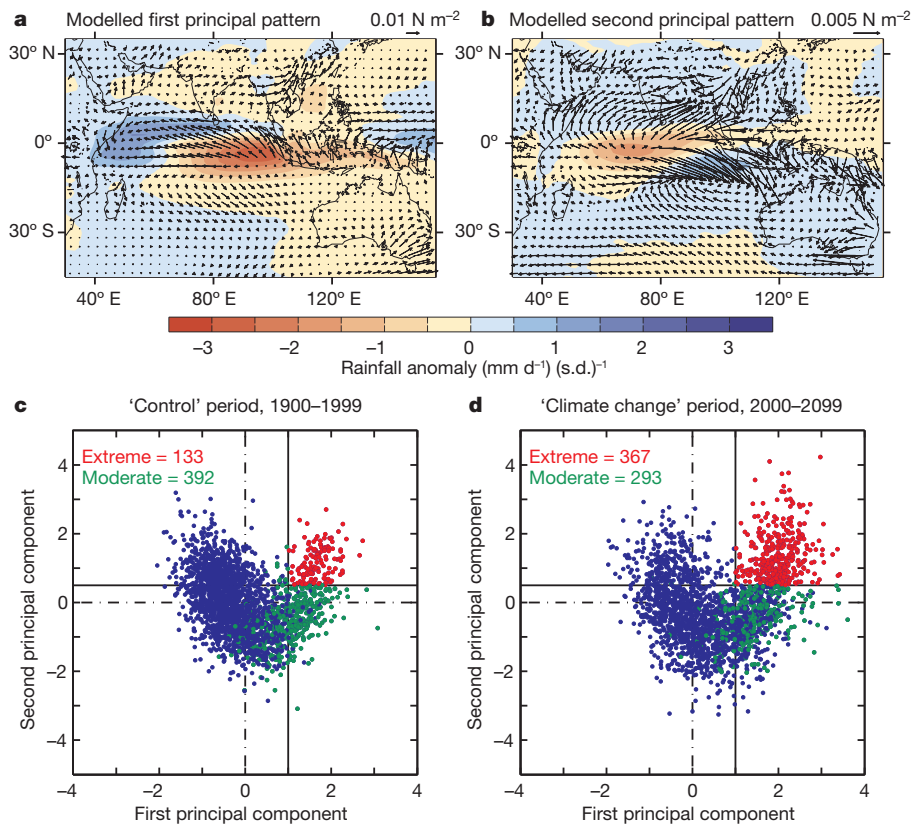
in the eastern pole, rendering a negative skewness of SST, in that cool anomalies off Sumatra grow to greater amplitude than warm anomalies<sup>25</sup>.

The strong nonlinearity along the Equator means that the representation of extreme pIOD impacts requires more than just the commonly used DMI. This along-the-Equator nonlinearity can be represented by two modes of empirical orthogonal function (EOF) of rainfall anomalies. The pattern of the first EOF (EOF1, 43.4% of the total variance, Fig. 1c) shows an east–west dipole of reduced convection, featuring anomalously cold SSTs and a shallow thermocline in the east but anomalies of opposite polarities in the west (Extended Data Fig. 3). This reflects characteristics

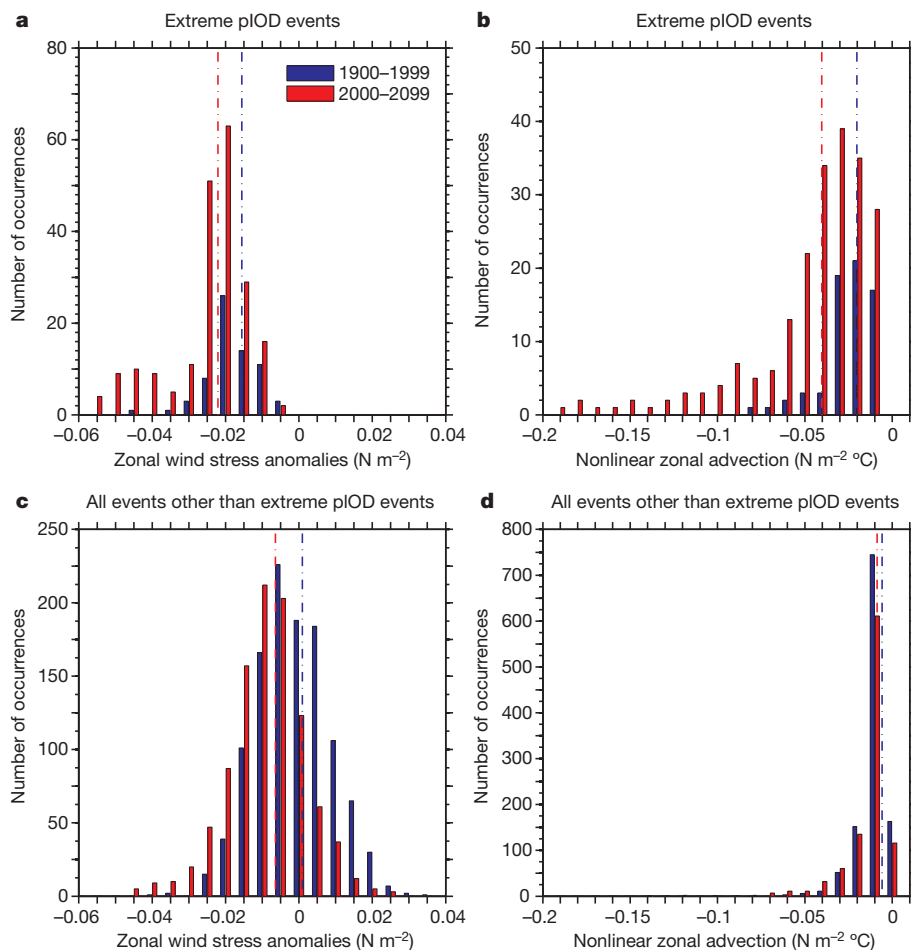
of pIOD events commonly depicted by the DMI. EOF2, which accounts for 20.7% of the total variance, on the other hand, reflects pronounced anomalous conditions during extreme pIOD events, as described above. Both EOFs feature enhanced convection over the western tropical Indian Ocean and equatorial Africa (Fig. 1d, Extended Data Fig. 3).

EOF1 and EOF2 (or the DMI) display a nonlinear relationship (Fig. 1e, f). During a moderate event, the two EOFs are of opposite sign. Thus, the associated rainfall anomalies tend to offset over the central Indian Ocean. In contrast, during an extreme pIOD, both EOFs are positive, rendering large negative rainfall anomalies that extend westward along

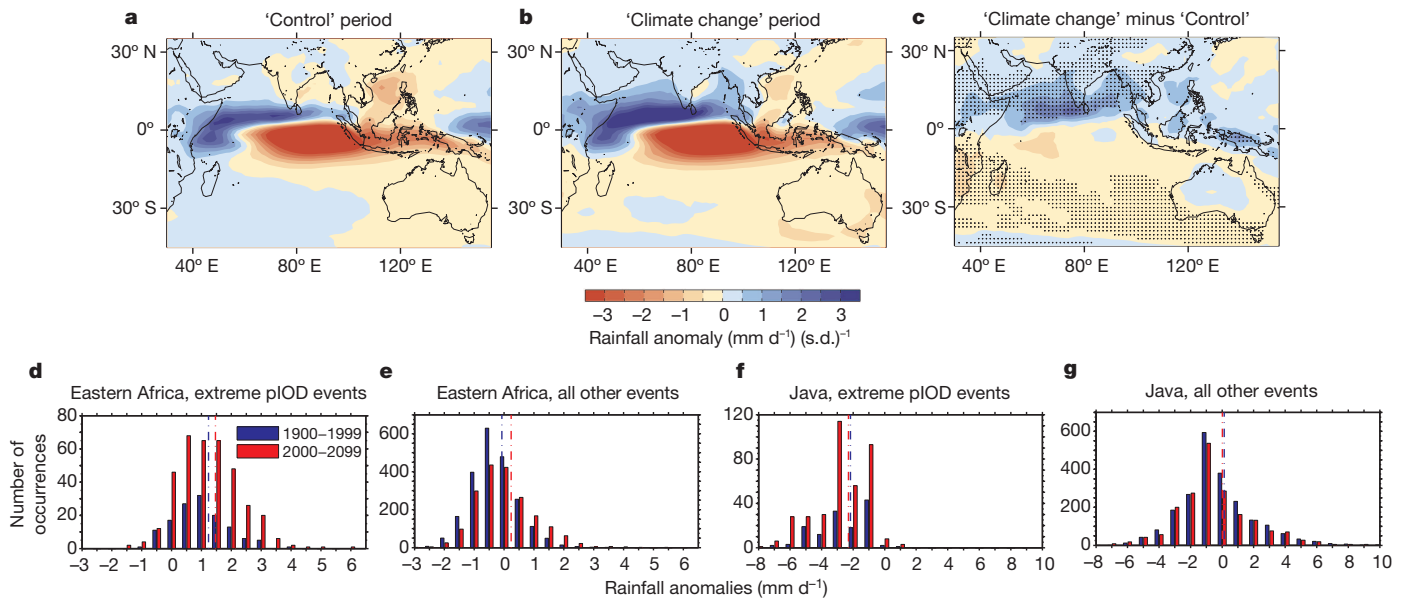




**Figure 2 | Multi-model ensemble average of the principal variability patterns of austral spring season rainfall and their nonlinear relationship.** **a, b**, First and second principal variability patterns of rainfall anomalies referenced to the 'control' period (1900–1999), obtained by applying an EOF analysis to rainfall anomalies in the equatorial region ( $10^{\circ} \text{S}$ – $10^{\circ} \text{N}$ ,  $40^{\circ} \text{E}$ – $100^{\circ} \text{E}$ ). Note the different vector scales in **a** and **b**. The associated pattern and wind stress vectors beyond the domain are obtained by a linear regression onto the EOF time series. The colour scale indicates rainfall in  $\text{mm day}^{-1}$  per 1.0 s.d. change; blue or red contours indicate increased or decreased rainfall. **c, d**, A nonlinear relationship between the principal component time series for the 'control' (1900–1999) and 'climate change' (2000–2099) periods. An extreme pIOD event (red dots) is defined as when the first principal component is greater than 1 s.d. and the second principal component is greater than 0.5 s.d. A moderate pIOD event (green dots) is determined from a detrended DMI when its amplitude is greater than 0.75 s.d. but is not an extreme pIOD event. Negative IOD and neutral years are indicated with blue dots. The number of extreme and moderate pIOD events is indicated.



**Figure 3 | Multi-model statistics associated with the increase in frequency of extreme pIOD events.** **a**, Multi-model ensemble histogram of zonal wind stress  $\tau_x$  anomalies in the equatorial Indian Ocean ( $5^{\circ} \text{S}$ – $5^{\circ} \text{N}$ ,  $60^{\circ} \text{E}$ – $100^{\circ} \text{E}$ ), referenced to the 'control' period. These are averaged over the July–October months of Indian Ocean dipole development phase. Values during extreme pIOD years in each period are separated into  $5 \times 10^{-3} \text{ N m}^{-2}$  bins centred at the tick point for the 'control' (blue) and 'climate change' (red) periods. The multi-model median for the 'control' (dashed blue line) and the 'climate change' (dashed red line) periods are indicated. **b**, The same as **a** but for the product of  $\tau_x$  anomalies shown in **a** multiplied by the DMI<sup>1</sup> (separated into  $0.01 \text{ N m}^{-2} \text{ } ^{\circ}\text{C}$  bins), approximating the nonlinear zonal advection (see Methods). **c, d**, The same as **a** and **b** but for all years excluding extreme pIOD events. The histogram for extreme pIOD is statistically different above the 95% confidence level from that for non-extreme pIOD events, for both the 'control' and the 'climate change' periods. On average, nonlinear advection is greater for extreme pIOD events than for non-extreme pIOD events.



**Figure 4 | Multi-model ensemble average of rainfall anomalies (referenced to the 'control' period) during extreme pIOD events and associated statistics in affected regions. a, b and c, Ensemble average rainfall anomalies in the 'control' and 'climate change' periods, and their difference ('climate change' minus 'control'). Stippling in c indicates regions where the differences are statistically significant at the 95% level as determined by a two-sided Student's *t*-test. d, e, Multi-model ensemble histogram of the rainfall anomalies over northern equatorial East Africa (0°–5°N, 37.5°E–47°E) for extreme pIOD events and all events other than extreme pIOD events, respectively. All extreme pIOD events**

in each period are separated into 0.5 mm day<sup>−1</sup> bins centred at the tick point for the 'control' (blue) and 'climate change' (red) periods. The multi-model mean for the 'control' (blue dashed line) and the 'climate change' (red dashed line) periods are indicated. In each period the histograms for extreme pIOD (for example, red bars in Fig. 4d) and non-extreme pIOD (for example, red bars in Fig. 4e) are statistically different above the 95% confidence level. f, g, The same as d and e, but for the Java region (8°S–6°S, 105.5°E–108.5°E), separated into 1 mm day<sup>−1</sup> bins. The two histograms in d and e are statistically different above the 95% confidence level, but this is not the case for the two histograms in f and g.

the Equator. The pIOD events of 1961, 1994 and 1997 are determined to be 'extreme' when EOF1 is greater than a 1-standard-deviation (s.d.) value and EOF2 is greater than a positive 0.5-s.d. value. The characteristics of the pIOD events are only fully captured by the superimposition of these two EOFs (Extended Data Fig. 4). Without EOF2, the salient feature of the westward extending equatorial anomalies that characterizes extreme pIOD would be missed. A similar EOF analysis on vertical velocity  $\omega$  at 500 mb, a measure of convection, generates similar patterns (Extended Data Fig. 5).

To assess the influence of greenhouse warming, we use the Coupled Model Intercomparison Project phase 5 (CMIP5<sup>26</sup>) multi-model database. The coupled general circulation models (CGCMs) used in this study are forced with historical anthropogenic and natural forcings, and future greenhouse-gas emission scenarios of Representative Concentration Pathway (RCP) 8.5, covering the 1900–2099 period. Not all of the 31 CGCMs considered here are able to simulate the characteristics of observed pIOD events. We focus on 23 CGCMs that simulate negative skewness of SST off Sumatra as well as the nonlinear relationship of the two rainfall EOFs (Extended Data Table 1, Fig. 2). An identical EOF analysis of  $\omega$  at 500 mb in 21 out of the 23 selected CGCMs, in which  $\omega$  is available, produces similar spatial patterns and their nonlinear relationship (Extended Data Fig. 6). From these 23 CGCMs, we define extreme pIOD events in the same manner as for the observed events, and compare their frequency in the first (1900–1999) and second (2000–2099) hundred-year periods. These two adjacent periods within a transient scenario are referred to as the 'control' and 'climate change' periods, respectively.

In aggregation, the frequency of extreme pIOD events based on rainfall EOFs increases by a factor of 2.7, from about one event every 17.3 years (133 events in 2,300 years) in the 'control' period, to one every 6.3 years (367 events in 2,300 years) in the 'climate change' period (Fig. 2c and d). This is statistically significant according to a bootstrap test<sup>27</sup>, underscored by a strong inter-model consensus, with 21 out of 23 models simulating an increase (Extended Data Table 1). Sensitivity tests to varying definitions of extreme pIOD further support the robustness of this result (Supplementary Tables 1 and 2).

Development of pIOD events can interact with an El Niño event<sup>28–30</sup>. The 1997 extreme pIOD developed in conjunction with the strongest El Niño of the twentieth century. The 1961 and 1994 extreme pIODs on the other hand occurred without an El Niño, supporting the notion that the generating mechanism behind an extreme pIOD event lies within the Indian Ocean<sup>2</sup>. We find no evidence that the increase in extreme pIOD events in the 'climate change' period is induced by a change in the frequency of El Niño or El Niño Modoki occurrences (see Methods and Extended Data Fig. 7).

Rather, the increase in extreme pIOD events appears to arise from mean state changes within the Indian Ocean (Extended Data Fig. 8), consistent with a weakening Walker circulation as projected under greenhouse warming<sup>19–21</sup>. Relative to the 'control' period, the altered mean state is more conducive to equatorial easterly winds, westward oceanic currents, an enhanced west-minus-east SST gradient, and the associated nonlinear zonal advection. There is a strong link between climatologically stronger easterly winds along the Equator and more occurrences of a given nonlinear advection (correlation coefficient  $r = 0.9$ , not shown). These changes lead to increasing occurrences of extreme pIOD events, because a smaller perturbation is required in the 'climate change' period to generate the same size of nonlinear zonal advection as seen during extreme pIOD events in the 'control' period (see Extended Data Fig. 9). Thus, there are increased occurrences of extreme pIOD events for a given size of the equatorial easterly anomaly (Fig. 3a), or a given strength of nonlinear advection (Fig. 3b). On the other hand, the changes associated with non-extreme pIOD events are not as apparent (Fig. 3c, d).

The increased frequency in extreme pIODs does not translate to greater intensity of rainfall anomalies over all regions affected by the pIOD (Fig. 4a–c). Over northeastern equatorial Africa, the extreme pIOD-induced wet events do become more intense in the 'climate change' period than in the 'control' period (Fig. 4d; the means are statistically different above the 95% confidence level). In contrast, there is no statistically significant difference between the two periods in the intensity of dry episodes over Java (Fig. 4f). In addition, the difference in rainfall intensity of the extreme events is generally

smaller than the difference in the mean rainfall (comparing Fig. 4a and Extended Data Fig. 8a), despite the far greater anomalies during extreme pIOD events. This illustrates that in general the impacts of extreme pIOD events experienced in the 'control' period will repeat more frequently in the 'climate change' period.

In summary, our finding of a greenhouse-induced increased frequency of extreme pIOD events is in stark contrast with previous results of no change in pIOD frequency about the evolving background state. By identifying nonlinear processes that give rise to extreme pIOD events, we show that under greenhouse warming, the evolving equatorial Indian Ocean towards climatologically stronger west-minus-east temperature gradients and easterly winds is more susceptible to producing more frequent extreme pIOD events. With the projected large increase in extreme pIOD events, we should expect more frequent occurrences of devastating weather events in affected regions.

## METHODS SUMMARY

The extreme pIOD events were diagnosed using a suite of distinct process-based indicators—such as anomalous equatorial easterlies, low rainfall and atmospheric subsidence—as induced by a downstream extension of the southeasterly trades. For observations, we focus on historical events in the satellite era (1979–present) monthly precipitation analysis, SSTs and other circulation fields from a global reanalysis (see Methods). We focus on austral spring, September–November, in which a pIOD typically peaks. A heat budget analysis for the eastern-to-central equatorial Indian Ocean using the European Centre for Medium-Range Weather Forecasts - Ocean Reanalysis System 3 reveals that the strong nonlinear zonal advection of heat sets the observed 1994 and 1997 extreme pIOD events apart from other events. The nonlinearity suggests that the traditional DMI, defined as the difference in SST anomalies between the western ( $50^{\circ}\text{E}$ – $70^{\circ}\text{E}$  and  $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ ) and eastern ( $90^{\circ}\text{E}$ – $110^{\circ}\text{E}$  and  $10^{\circ}\text{S}$ – $0^{\circ}\text{S}$ ) parts of the Indian Ocean<sup>1</sup> is not sufficient to differentiate an extreme pIOD event. Thus, we propose an identification method for extreme pIOD, in which we apply EOF analysis to rainfall anomalies and vertical velocity  $\omega$  at 500 mb in the equatorial Indian Ocean ( $40^{\circ}\text{E}$ – $100^{\circ}\text{E}$ ,  $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ ). This produces two principal variability patterns, one depicting an east–west pattern and the other depicting dry conditions along the central Indian Ocean extending from the east. An extreme pIOD event is defined as when the first principal time series is greater than 1 s.d., and the second greater than 0.5 s.d. This definition exclusively captures the three observed extreme pIOD events. To select CGCMs, the method is applied to 31 CMIP5 CGCMs, each covering 105 years of a pre-twenty-first-century climate change simulation using historical anthropogenic and natural forcings (1901–2005) and a further 95 years (2006–2100) under the RCP8.5 forcing scenario<sup>26</sup>.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 28 November 2013; accepted 8 April 2014.**

1. Saji, N. H., Goswami, B. N., Vinayachandran, P. N. & Yamagata, T. A dipole in the tropical Indian Ocean. *Nature* **401**, 360–363 (1999).
2. Webster, P. J., Moore, A. M., Loschnigg, J. P. & Leben, R. R. Coupled oceanic-atmospheric dynamics in the Indian Ocean during 1997–98. *Nature* **401**, 356–360 (1999).
3. Yu, L. & Rienecker, M. M. Mechanisms for the Indian Ocean warming during the 1997–98 El Niño. *Geophys. Res. Lett.* **26**, 735–738 (1999).
4. Murtugudde, R., McCreary, J. P. & Busalacchi, A. J. Oceanic processes associated with anomalous events in the Indian Ocean with relevance to 1997–1998. *J. Geophys. Res.* **105**, 3295–3306 (2000).
5. Meyers, G. A., McIntosh, P. C., Pigot, L. & Pook, M. J. The years of El Niño, La Niña, and interactions with the tropical Indian Ocean. *J. Clim.* **20**, 2872–2880 (2007).

6. Ummenhofer, C. C. *et al.* What causes southeast Australia's worst droughts? *Geophys. Res. Lett.* **36**, L04706 (2009).
7. Ashok, K., Guan, Z. & Yamagata, T. Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophys. Res. Lett.* **30**, 1821 (2003).
8. Cai, W., Cowan, T. & Raupach, M. Positive Indian Ocean Dipole events precondition southeast Australia bushfires. *Geophys. Res. Lett.* **36**, L19710 (2009).
9. Emmanuel, S. Impact to lung health of haze from forest fires: the Singapore experience. *Respirology* **5**, 175–182 (2000).
10. Frankenberg, E., McKee, D. & Thomas, D. Health consequences of forest fires in Indonesia. *Demography* **42**, 109–129 (2005).
11. Zubair, L., Rao, S. A. & Yamagata, T. Modulation of Sri Lankan Maha rainfall by the Indian Ocean dipole. *Geophys. Res. Lett.* **30**, 1063 (2003).
12. Abram, N. J., Gagan, M. K., McCulloch, M. T., Chappell, J. & Hantoro, W. S. Coral reef death during the 1997 Indian Ocean Dipole linked to Indonesian wildfires. *Science* **301**, 952–955 (2003).
13. Behera, S. K. *et al.* Paramount impact of the Indian Ocean Dipole on the East African short rains: a CGCM study. *J. Clim.* **18**, 4514–4530 (2005).
14. Black, E., Slingo, J. & Sperber, K. R. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Mon. Weath. Rev.* **131**, 74–94 (2003).
15. Hashizume, M., Chaves, L. F. & Minakawa, N. Indian Ocean Dipole drives malaria resurgence in East African highlands. *Sci. Rep.* **2**, 269 (2012).
16. Schott, F. A., Xie, S.-P. & McCreary, J. P. Indian Ocean circulation and climate variability. *Rev. Geophys.* **47**, RG1002 (2009).
17. Page, S. E. *et al.* The amount of carbon released from peat and forest fires in Indonesia in 1997. *Nature* **420**, 61–65 (2002).
18. Udea, H. & Matsumoto, J. A. Possible triggering process of east–west asymmetric anomalies over the Indian Ocean in relation to 1997/1998 El Niño. *J. Meteorol. Soc. Jpn* **78**, 803–818 (2000).
19. Vecchi, G. A. & Soden, B. J. Global warming and the weakening of the tropical circulation. *J. Clim.* **20**, 4316–4340 (2007).
20. Zheng, X. T. *et al.* Indian Ocean Dipole response to global warming in the CMIP5 multimodel ensemble. *J. Clim.* **26**, 6067–6080 (2013).
21. Cai, W. *et al.* Projected response of the Indian Ocean Dipole to greenhouse warming. *Nature Geosci.* **6**, 999–1007 (2013).
22. Wijffels, S. & Meyers, G. An intersection of oceanic waveguides: variability in the Indonesian throughflow region. *J. Phys. Oceanogr.* **34**, 1232–1253 (2004).
23. Feng, M. & Meyers, G. Interannual variability in the tropical Indian Ocean: a two-year time scale of Indian Ocean Dipole. *Deep-Sea Res.* **50**, 2263–2284 (2003).
24. Wyrtki, K. An equatorial jet in the Indian Ocean. *Science* **181**, 262–264 (1973).
25. Hong, C. C., Li, T., Lin, H. & Kug, J. S. Asymmetry of the Indian Ocean Dipole. Part I: Observational analysis. *J. Clim.* **21**, 4834–4848 (2008).
26. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experimental design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
27. Austin, P. C. Bootstrap methods for developing predictive models. *Am. Stat.* **58**, 131–137 (2004).
28. Fischer, A. *et al.* Two independent triggers for the Indian Ocean dipole/zonal mode in a coupled GCM. *J. Clim.* **18**, 3428–3449 (2005).
29. Luo, J. J. *et al.* Interaction between El Niño and extreme Indian Ocean Dipole. *J. Clim.* **23**, 726–742 (2010).
30. Izumo, T. *et al.* Influence of the state of the Indian Ocean Dipole on the following year's El Niño. *Nature Geosci.* **3**, 168–172 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** W.C. and E.W. are supported by the Australian Climate Change Science Program, and the Goyder Institute. W.C. is also supported by a CSIRO Office of Chief Executive Science Leader award. A.S. is supported by the Australian Research Council.

**Author Contributions** W.C. conceived the study and directed the analysis. G.W. and E.W. performed the model output analysis. A.S. conducted the heat budget analysis. W.C. wrote the initial draft of the paper. All authors contributed to interpreting results, discussion of the associated dynamics and improvement of this paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.C. ([wenju.cai@csiro.au](mailto:wenju.cai@csiro.au)).



## METHODS

**Data, reanalyses and EOF analysis.** We used data in the satellite era (1979–present) which include Global Precipitation Climatology Project monthly precipitation analysis<sup>31</sup>, global analyses of SSTs<sup>32</sup>, and circulation fields from the National Center for Environmental Prediction and National Center for Atmospheric Research global reanalysis<sup>33</sup>. Ocean column data of velocities and temperatures for heat budget analysis are based on the European Centre for Medium-Range Weather Forecasts - Ocean Reanalysis System 3 (ECMWF ORA-S3)<sup>34</sup>. We use a multivariate signal processing method referred to as EOF analysis<sup>35</sup> to anomalies of rainfall and vertical velocity  $\omega$  at 500 mb (ref. 33). The EOF technique deconvolves the spatio-temporal variability into orthogonal modes, each described by a principal spatial pattern and an associated principal component time series.

**Heat budget analysis.** We examine the surface heat balance of the tropical Indian Ocean which can be expressed as:

$$\begin{aligned} \partial T^a / \partial t = & -[(u^a \partial T^a / \partial x + \bar{u} \partial T^a / \partial x + u^a \partial \bar{T} / \partial x) \\ & + (v^a \partial T^a / \partial y + \bar{v} \partial T^a / \partial y + v^a \partial \bar{T} / \partial y) \\ & + (w^a \partial T^a / \partial z + \bar{w} \partial T^a / \partial z + w^a \partial \bar{T} / \partial z)] + Q + \text{residual} \end{aligned} \quad (1)$$

The variables  $T$ ,  $u$ ,  $v$  and  $w$  are potential temperature, and the zonal, meridional and vertical ocean current velocities, respectively, averaged over the top 50 m. Differential operators,  $x$ ,  $y$ ,  $z$  and  $t$ , are along the zonal, meridional and vertical directions, and time, respectively. All variables are derived from the ECMWF ORA-S3 observational data assimilation system<sup>34</sup> at a horizontal resolution of  $1^\circ$  latitude by  $1^\circ$  longitude, increasing to  $0.3^\circ$  in latitude towards the Equator. The rate of change of the mixed layer temperature ( $dT/dt$ ) is calculated using a centred-difference approximation. Superscript 'a' and overbar denote anomalous and long-term averaged quantities, respectively. Equation (1) states that the rate of change or tendency of the surface temperature is balanced by zonal advection of heat by the zonal current (first bracketed terms on the right hand side), meridional advection (second bracketed terms), vertical advection (third bracketed terms), the net surface air–sea heat flux ( $Q$ ), and all other factors not explicitly expressed (residual), such as mixing and diffusion.

We use the entire reanalysis period of the ORA-S3, which spans 1959–2006, to examine processes during the 1961 event. All variables in equation (1) are linearly detrended and averaged over the eastern-to-central equatorial region between  $5^\circ$  S– $5^\circ$  N and  $60^\circ$  E– $100^\circ$  E, over which the 1961, 1994 and 1997 extreme events emerge as the only pIODs that exhibit large anomalous cooling (Extended Data Fig. 1a). We examine the heat budget terms averaged over the developing period of Indian Ocean dipole events (July–October; Extended Data Fig. 1b). It may be noted that the 1997 event exhibits exceptionally strong and prolonged cooling compared to the 1961 and 1994 events, which see an earlier start of the cooling at the end of boreal spring.

The nonlinear vertical advective process ( $w^a \partial T^a / \partial z$ ), that is, the process associated with anomalous upwelling and anomalous vertical temperature gradients, contributes substantially to the cooling of the equatorial Pacific during these events, especially during moderate pIOD events (Extended Data Fig. 1c). During the 1961, 1994 and 1997 events, however, the nonlinear zonal advection term ( $u^a \partial T^a / \partial x$ ) is exceptionally strong, extending notably farther to the west from the eastern Indian Ocean, as compared to the other events (Extended Data Fig. 2a–d). Although the nonlinear vertical advection is more prominent during the 1961 and 1997 events (Extended Data Fig. 1c), which is in part also driven by the equatorial easterly winds, it is the nonlinear zonal advection that sets the 1961, 1994 and 1997 events apart from the rest of the pIODs. This stems from the exceptionally strong westward current and its associated easterly winds (Extended Data Fig. 2f).

As shown in Extended Data Fig. 2f, the significant correlation between the zonal wind and current ( $r = 0.87$ ) means that the nonlinear zonal advection term over the equatorial region can be well approximated as a product between the zonal wind (averaged over  $5^\circ$  S– $5^\circ$  N,  $60^\circ$  E– $100^\circ$  E) and the DMI<sup>3</sup>:

$$-u^a \partial T^a / \partial x \approx \tau^x \text{DMI} \alpha \quad (2)$$

with  $\alpha = \beta/L$ , where  $\beta$  is the regression coefficient between the zonal wind and zonal current, and  $L$  is the longitudinal width of the equatorial box (Extended Data Fig. 2a). This parameterization is used to represent the nonlinear zonal advection term in the CGCMs.

Strikingly similar to the nonlinear zonal advection, the proxy exclusively identifies the three observed extreme events. It may be noted that the proxy is further from the actual value for the 1961 and 1997 events than for the 1994 event. This is expected, owing to the particularly strong nonlinear vertical advection of the 1961 and 1997 events. Using 93-year time series of nine CMIP5 models for which we had access to the required variables, the robustness of the proxy is signified by the high positive correlation coefficient with the nonlinear advection, ranging from 0.59 to 0.92, significant above the 99% confidence level.

**Characterization of extreme pIOD events.** The extreme pIOD events were characterized using a suite of distinctive process-based indicators, such as anomalous

equatorial easterlies, low rainfall and atmospheric subsidence as induced by a downstream extension of the southeasterly trade winds. For observations, we focus on historical events in the satellite era (1979–present) using Global Precipitation Climatology Project monthly precipitation analysis<sup>31</sup> from <http://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html> and SSTs<sup>32</sup> and other circulation fields from a global reanalysis<sup>33</sup>. We focus on austral spring, September–November, in which a pIOD typically peaks, and apply EOF analysis<sup>35</sup> to rainfall anomalies and  $\omega$  (ref. 33) at 500 mb in the equatorial Indian Ocean ( $40^\circ$  E– $100^\circ$  E,  $10^\circ$  S– $10^\circ$  N). This produces two principal variability patterns. The first principal pattern reflects a strong rainfall reduction over the eastern pole accompanied by a moderately rainfall increase over the western equatorial Indian Ocean, and the second principal pattern is characterized by a westward extension of rainfall reduction from the east, accompanied by a rainfall increase farther west near equatorial east Africa. Note that the wetting and northwesterly winds off Sumatra in EOF2 are opposite to the drying and southeasterly anomalies in EOF1. Along the Equator, dry anomalies north of Sumatra embedded in EOF2 oppose weak wet anomalies in EOF1. The opposing polarity highlights that during extreme pIOD events, the centre of cold and dry anomalies is not concentrated in the Sumatra region but shifts northward for a westward extension along the Equator.

**Model selection.** We utilize 31 CMIP5 CGCMs forced with historical anthropogenic and natural forcings, and future greenhouse gas under emission scenario of Representative Concentration Pathway (RCP) 8.5<sup>26</sup>, covering a 200-year period. Two features of the nonlinearity associated with extreme pIODs are used to select models. These are the negative skewness of SST anomalies over the eastern pole, and the nonlinear positive feedback along the Equator involving the west-minus-east SST gradient, wind and oceanic currents, and nonlinear zonal advection, as indicated by a nonlinear relationship between the two EOFs. These two features are not mutually inclusive and are both used in our study.

Although the majority of CGCMs generate variability like that of the Indian Ocean dipole, only a subgroup of CGCMs simulate the observed nonlinear ocean–atmosphere coupling over the eastern Indian Ocean as depicted by the negative skewness of SST anomalies over the eastern pole during the austral spring (September–November), which is  $-0.85$  in observations since 1979. The level of nonlinearity varies vastly among CGCMs, and we consider negative skewness of any extent. Out of the 31 CGCMs, 23 satisfy the SST skewness criterion. The selected CGCMs yield a mean skewness of  $-0.84$ , close to the observed (Extended Data Table 1).

All selected 23 CGCMs reproduce the observed IOD pattern obtained by regressing September–November SST anomalies onto the DMI, with a pattern correlation greater than 0.75 (Supplementary Table 3). The same EOF analysis is carried out for each individual model using rainfall anomalies referenced to the mean over the 'control' period. Prior to the analysis, data are interpolated into a common grid of  $1.5^\circ$  latitude by  $1.5^\circ$  longitude. Our EOF outputs are scaled so that the EOF time series have a standard deviation of one to facilitate an inter-model comparison and aggregation. Details of the variance explained by EOF1 and EOF2 are listed in Supplementary Table 3. All 23 models produce the nonlinear relationship between the two leading rainfall EOFs, indicating their ability to generate the nonlinear equatorial positive feedback associated with the extreme pIOD. Outputs of  $\omega$  at 500 mb are available in 21 of the 23 CGCMs, and a nonlinear relationship between the two leading vertical velocity EOFs is generated in all the 21 models.

We derive changes in the occurrence of extreme pIOD events by comparing the frequency of the first 100 years ('control' period) to that of the second 100 years ('climate change' period). Of the eight CGCMs which are not able to simulate the negative SST skewness, only three CGCMs are not able to reproduce the nonlinear relationship between the two rainfall EOFs, suggesting that the negative skewness of SST anomalies in the eastern pole is not a prerequisite for the equatorial positive feedback associated with extreme pIOD events. We also test the sensitivity of our results to varying definitions (Supplementary Tables 1 and 2), including a case in which the criterion of the negative SST skewness is excluded: that is, including all 31 CGCMs. In all cases, there is a statistically significant increase (greater than a 130% increase) in the occurrences of extreme pIOD events from the 'control' to the 'climate change' period.

**Occurrences of extreme pIOD and the El Niño.** The modelled increase in extreme pIOD events is not induced by a change in the frequency in El Niño occurrences, because there is no inter-model consensus between the two periods in the frequency change of El Niño defined as when the quadratically detrended Niño3 ( $5^\circ$  S– $5^\circ$  N,  $150^\circ$  W– $90^\circ$  W) SST is greater than 0.5 s.d. (Extended Data Fig. 7a), consistent with previous studies<sup>36,37</sup>. Nor is there a statistically significant relationship between changes in the number of extreme pIOD events and changes in the number of El Niño events (Extended Data Fig. 7a), extreme El Niño defined as with Niño3 rainfall greater than a threshold value (Extended Data Fig. 7b)<sup>38</sup>, or detrended Niño3 SST greater than a threshold value (for example, 1.5 s.d.) (Extended Data Fig. 7c). In addition, there is no systematic change in the relationship between the Indian Ocean dipole and the El Niño/Southern Oscillation (ENSO) (Extended Data Fig. 7d)<sup>21</sup>. Similarly, there is no inter-model consensus on how Modoki El Niño, defined as occurring when the index<sup>39</sup>

is greater than a 0.5 s.d., will change. Nor is there a statistically significant relationship between changes in the number of extreme pIOD events and changes in the number of Modoki El Niño events (Extended Data Fig. 7e), and there is little change in the relationship between the Indian Ocean dipole and the Modoki ENSO (Extended Data Fig. 7f).

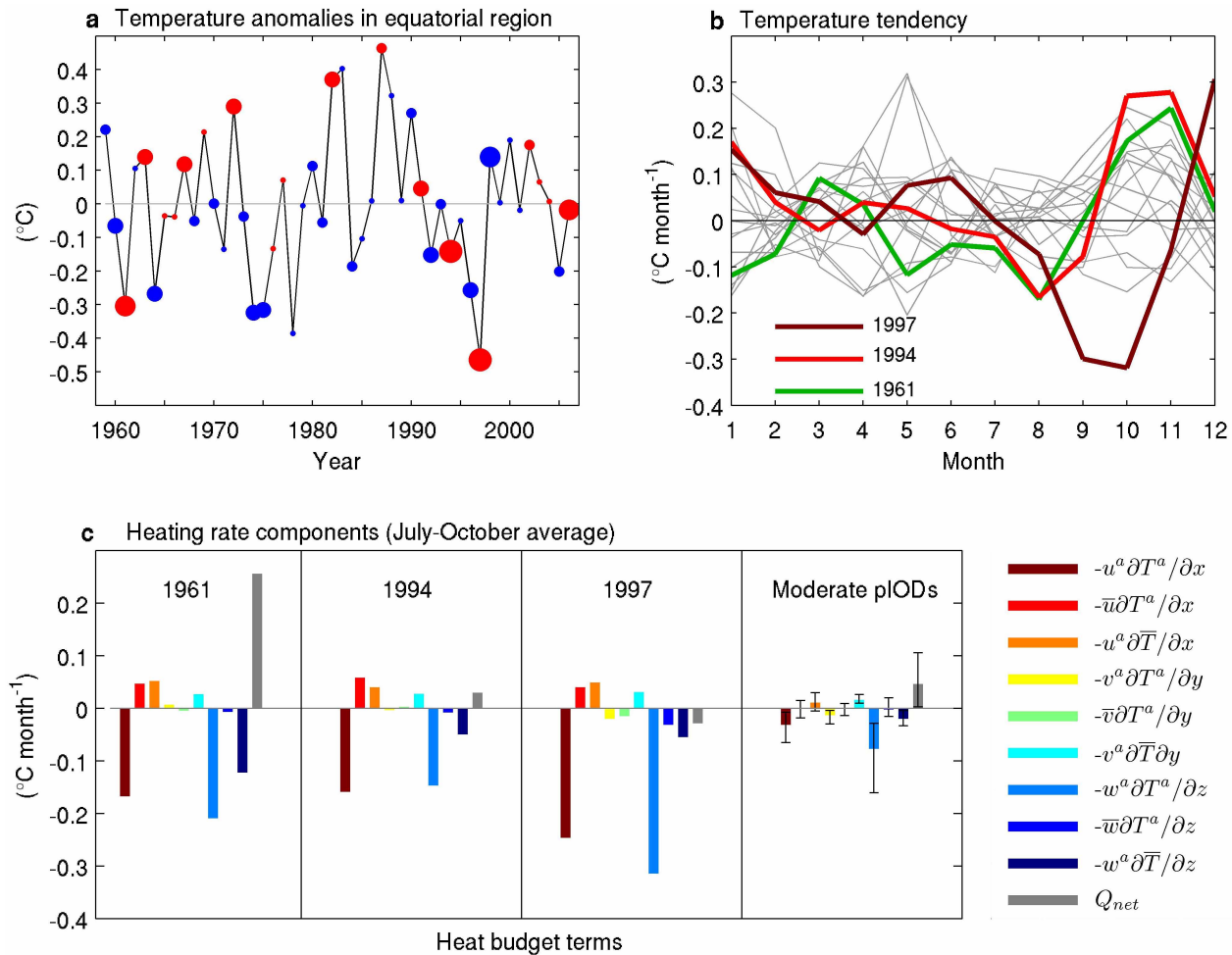
**Statistical significance test.** We use a bootstrap method<sup>27</sup> to examine whether the change in frequency of the extreme pIOD events is statistically significant. The 2,300 samples from the 23 CMIP5 CGCMs in the 'control' period are re-sampled randomly to construct another 10,000 realizations of 2,300-year records. In the random re-sampling process, any extreme pIOD event is allowed to be selected again. The standard deviation of the extreme pIOD frequency in the inter-realization is 11.2 events per 2,300 years, far smaller than the difference of 234 events per 2,300 years between the 'control' and the 'climate change' periods (Fig. 2c, d), indicating a strong statistical significance. The maximum frequency is 176, far smaller than that in the 'climate change' period of 367. Increasing the realizations to 20,000 or 30,000 yields essentially an identical result.

To further confirm the statistical significance of our result with ample samples of IOD behaviour across a longer time series without climate change forcing, we use a Canadian model (CanESM2), in which a pre-industrial simulation of 996 years. We examine the rarity of extreme pIOD event relative to that in the 'climate change' period with this same model. In the pre-industrial period the frequency is one per 13 years, but in the 'climate change' period there is a 180% increase to one event per 5 years. In the pre-industrial period, such an extreme pIOD event is far rarer.

Dividing the 996 years into 9 sets of 100 years and a set of 96 years, we find no frequency in these sets is as high as that in the 'climate change' period. The lowest

frequency is one event per 25 years, and the highest frequency, one event per 7.7 years, is 50% lower than the frequency in the 'climate change' period. This highlights the robustness of the greenhouse-warming-induced increase in the extreme pIOD frequency, above that generated by natural variability, which is represented by the spread of inter-model differences.

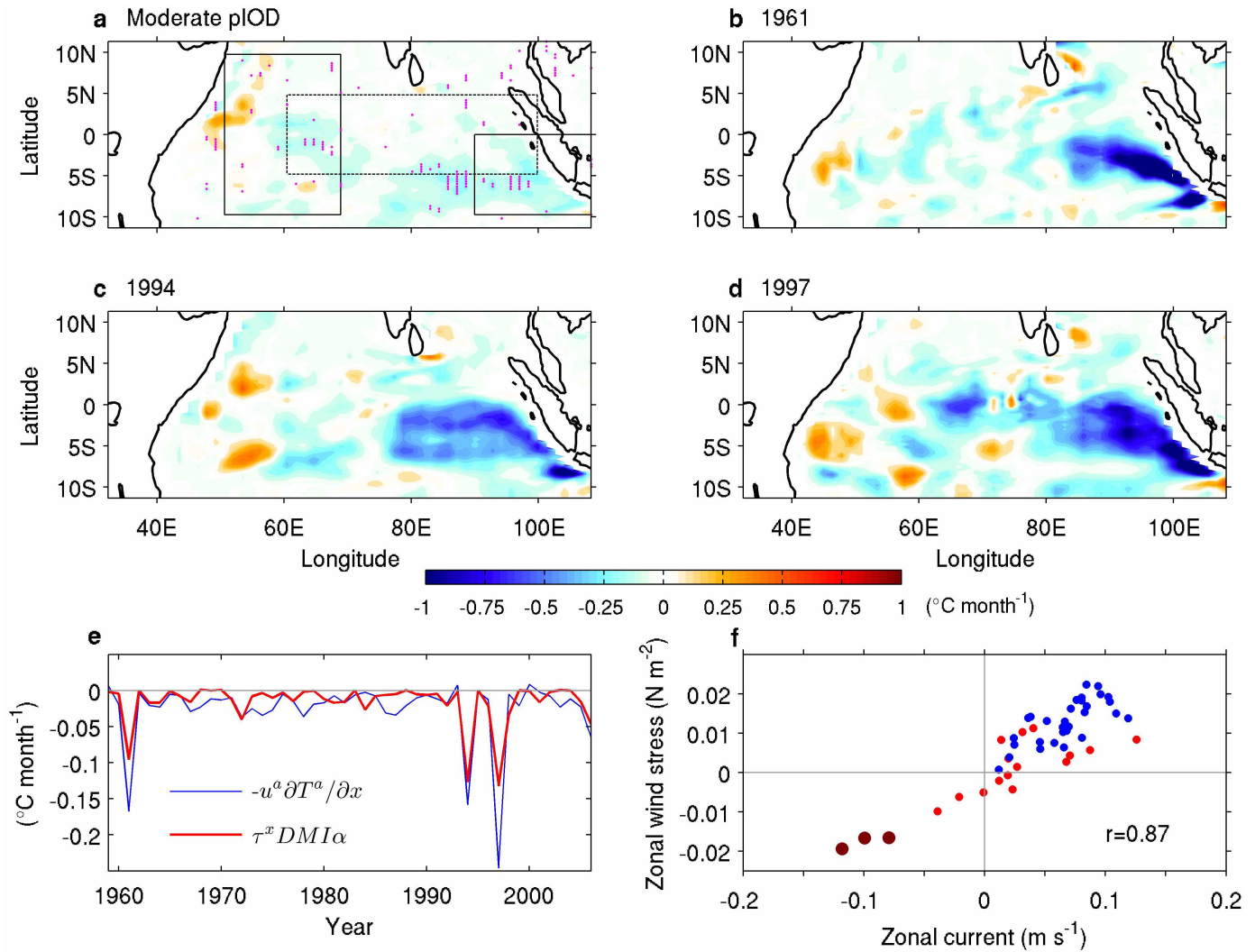
31. Adler, R. F. *et al.* The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4**, 1147–1167 (2003).
32. Rayner, N. A. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).
33. Kalnay, E. *et al.* The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
34. Balmaseda, M. A., Vidard, A. & Anderson, D. The ECMWF ocean analysis system: ORA-S3. *Mon. Weath. Rev.* **136**, 3018–3034 (2008).
35. Lorenz, E. N. *Empirical Orthogonal Functions and Statistical Weather Prediction*. Statistical Forecast Project Report 1 (MIT Department of Meteorology, 1956).
36. Guilyardi, E. *et al.* Understanding El Niño in ocean–atmosphere general circulation models: progress and challenges. *Bull. Am. Meteorol. Soc.* **90**, 325–340 (2009).
37. Collins, M. *et al.* The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geosci.* **3**, 391–397 (2010).
38. Cai, W. *et al.* Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Clim. Change* **4**, 111–116 (2014).
39. Ashok, K., Behera, S. K., Rao, S. A., Weng, H. & Yamagata, T. El Niño Modoki and its possible teleconnection. *J. Geophys. Res.* **112**, C11007 (2007).



**Extended Data Figure 1 | Heat budget analysis of the extreme pIOD events based on an ocean reanalysis<sup>34</sup>.** **a**, Temperature anomalies averaged over  $5^{\circ}\text{S}$ – $5^{\circ}\text{N}$  and  $60^{\circ}\text{E}$ – $100^{\circ}\text{E}$ , over the top 50 m, and over September–November. The filled blue and red circles indicate negative and positive DMI, with the size of the markers indicating the relative strength of the DMI. **b**, The rate of change of the temperature anomalies as a function of calendar month for all positive DMI values, with that of 1961 shown in green, 1994 in light red and

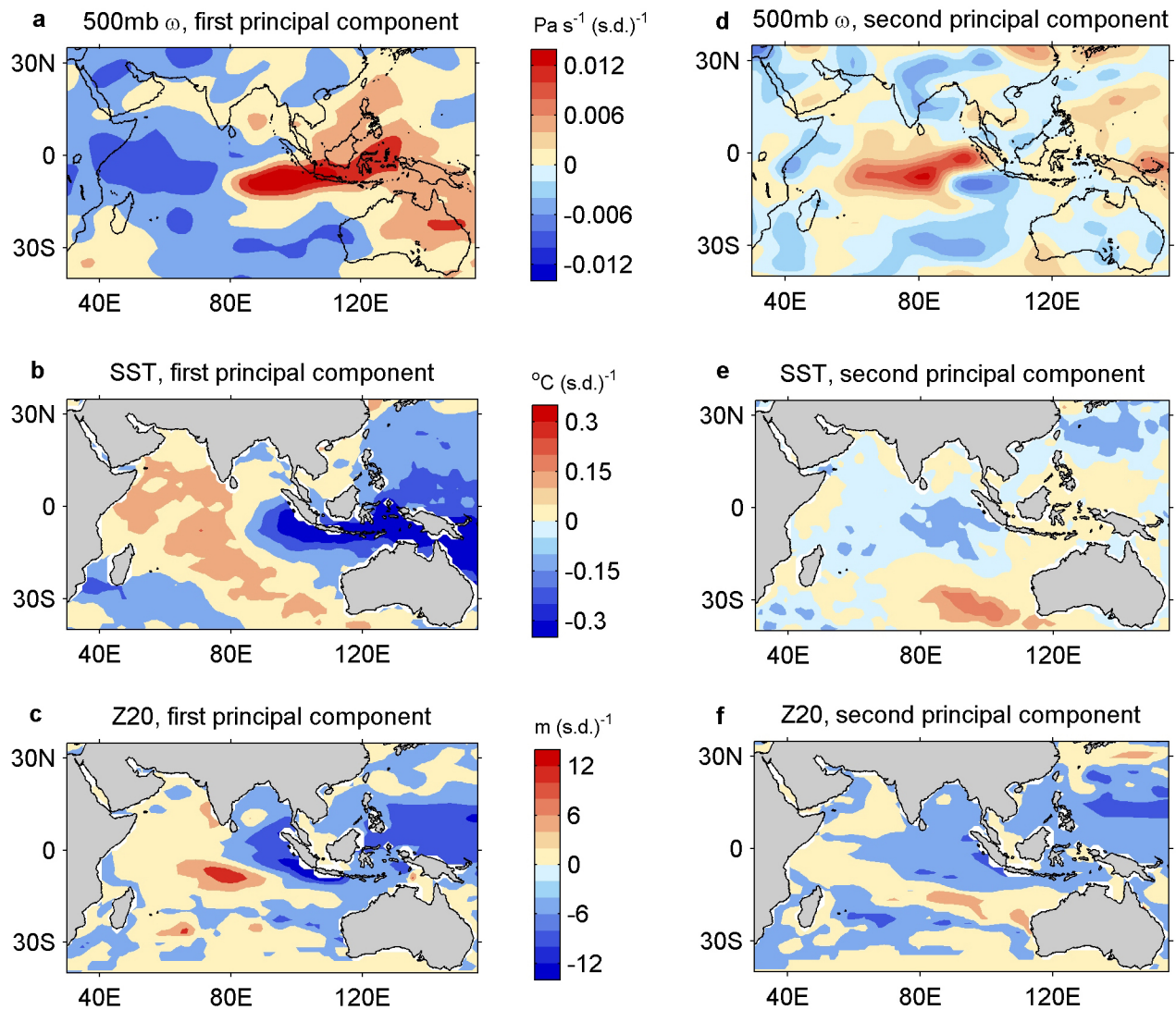
1997 in dark red, and all others in grey. **c**, The heat budget components averaged over July–October of Indian Ocean dipole development phase, for the 1961, 1994 and 1997 extreme events, and a composite of moderate pIOD events in the satellite era (1982, 1987, 2002 and 2006). The uncertainty bar on each composite represents the range of values over the four moderate pIOD events. The nonlinear zonal advection term ( $u^a \partial T^a / \partial x$ ) (dark red in **c**) is particularly large during the 1961, 1994 and 1997 events (see Methods for more details).





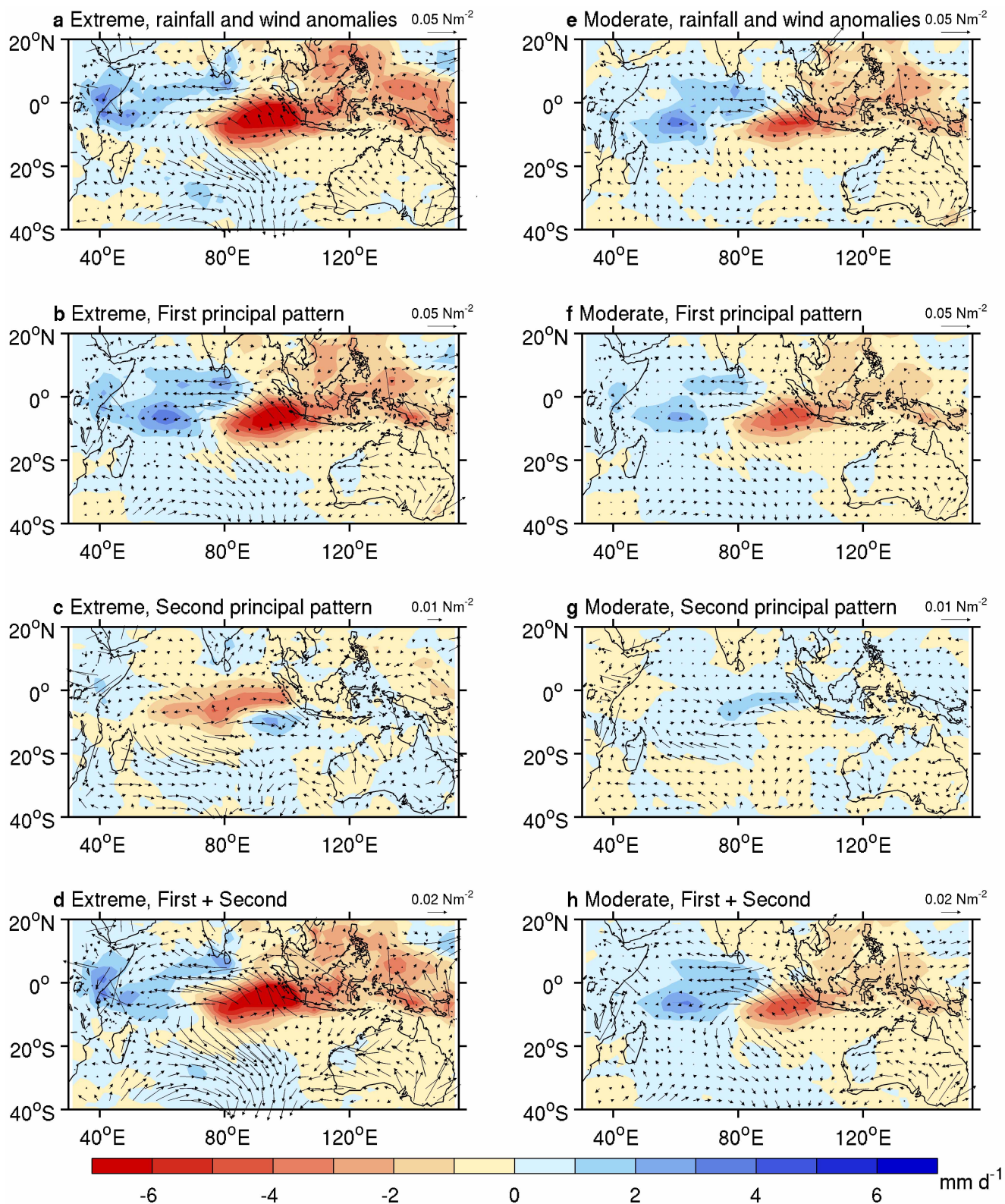
**Extended Data Figure 2 | Nonlinear zonal advection term over the growth phase of pIOD events.** The nonlinear zonal advection term ( $u^a \partial T^a / \partial x$ ) averaged over July to October for: **a**, a composite of moderate pIOD events, **b**, the 1961 pIOD event, **c**, the 1994 pIOD event and **d**, the 1997 pIOD event. The moderate pIOD events taken for the composite in **a** are the those in the satellite era: the 1982, 1987, 2002 and 2006 events. Stippled locations in **a** indicate composite values that are significant above the 95% confidence level ( $P$ -value  $< 0.05$ ) according to a Student's  $t$ -test. **e**, The approximation of the

nonlinear advection term,  $u^a \partial T^a / \partial x$ , averaged over the equatorial boxed region ( $5^{\circ}\text{S}$ – $5^{\circ}\text{N}$  and  $60^{\circ}\text{E}$ – $100^{\circ}\text{E}$ ; as shown in **a**) using the product between the corresponding zonal wind stress and the DMI (see Methods). The DMI is a measure of zonal gradient of temperature anomalies averaged over the western and eastern boxed regions in **a**, **f**. The total zonal current versus total zonal wind stress averaged over the equatorial box region in **a**. A particularly strong zonal current reversal is seen during the 1961, 1994 and 1997 pIOD events (large red dots in **f**, see Extended Data Fig. 1a).



**Extended Data Figure 3 | Circulation anomalies associated with the principal variability patterns of austral spring (September–November) rainfall.** **a–c**, Vertical velocity  $\omega$  at 500 mb ( $\text{Pa s}^{-1}$ ) from reanalysis data<sup>33</sup> (positive indicating descending motion) (**a**), SST ( $^{\circ}\text{C}$ ) (ref. 32) (**b**) and thermocline depth (m) (ref. 34) (**c**) anomalies associated with the first principal

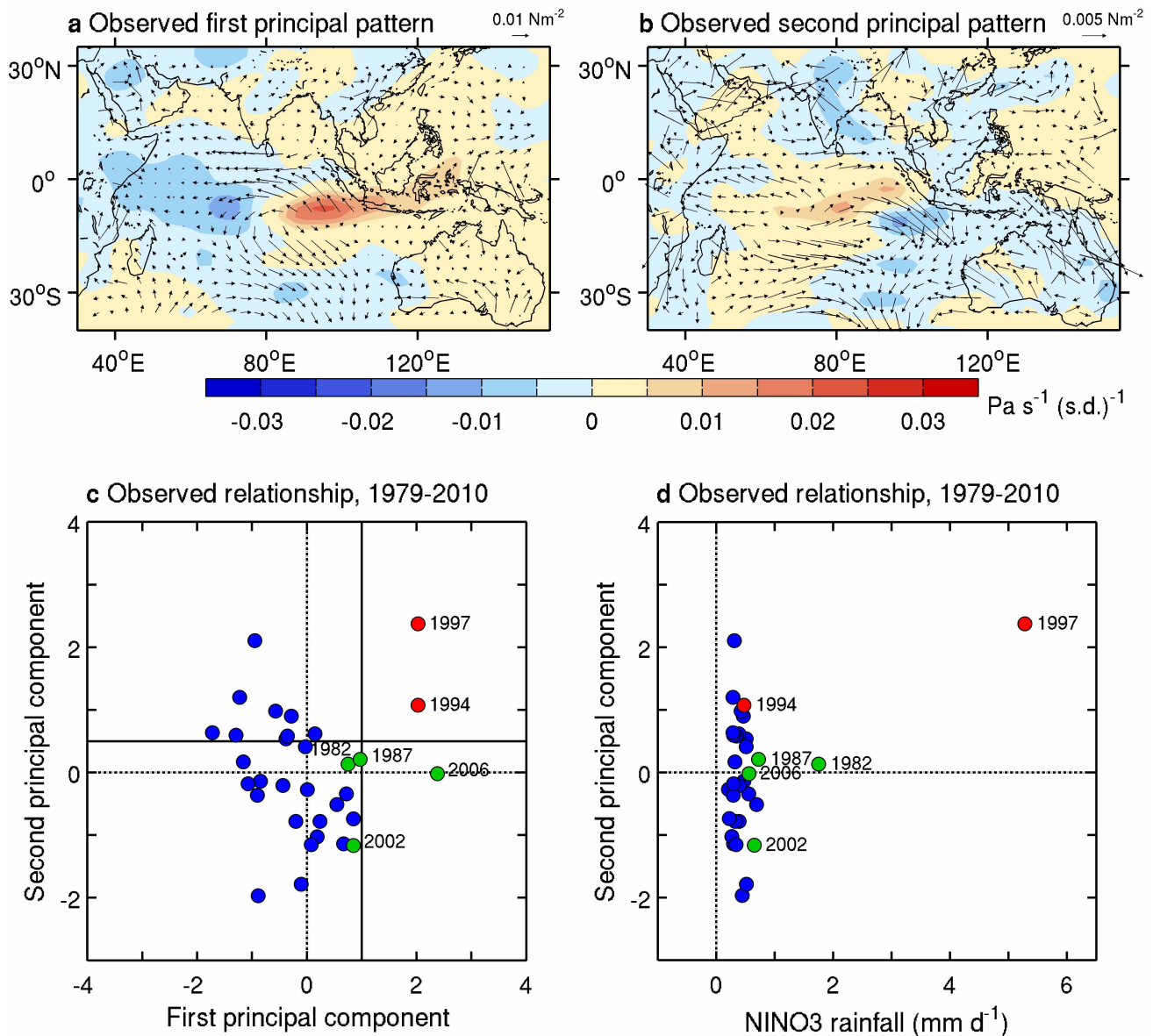
variability pattern (Fig. 1c). The patterns are obtained through linear regression of the corresponding variables onto the principal component time series of EOF1. **d–f**, The same as for **a–c**, but for the second principal variability pattern (Fig. 1d).



**Extended Data Figure 4 | Reconstruction of an extreme pIOD and a moderate pIOD event using the first two principal rainfall variability patterns.** a–d, Composite of anomalies associated with the 1994 and 1997 extreme pIOD events, showing the observed rainfall and wind stress anomalies, and anomalies reconstructed from the first principal, the second principal, and the first and second principal components combined, using satellite-era rainfall anomaly data from the Global Precipitation

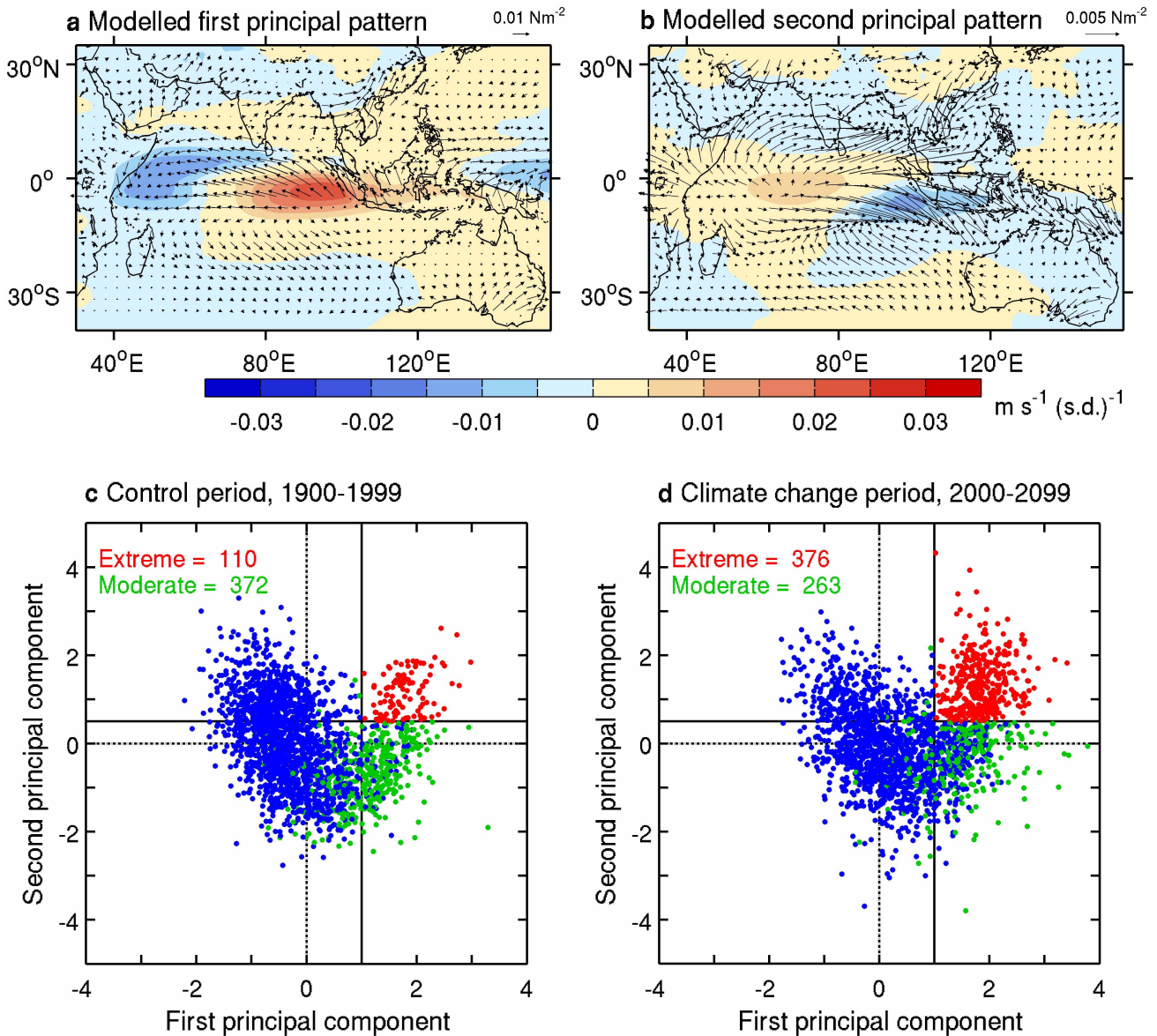
Climatology Project version 2 (ref. 31) and reanalysis wind stress<sup>33</sup>. Note the different vector scales shown in the top right corner for each panel. e–h, The same as a–d, but for composites of anomalies associated with the 1982, 1987, 2002 and 2006 moderate pIOD events. The exercise highlights that the difference between a moderate and an extreme pIOD depends on the role of the second principal component, and can only be realized with the use of both of the two principal components.





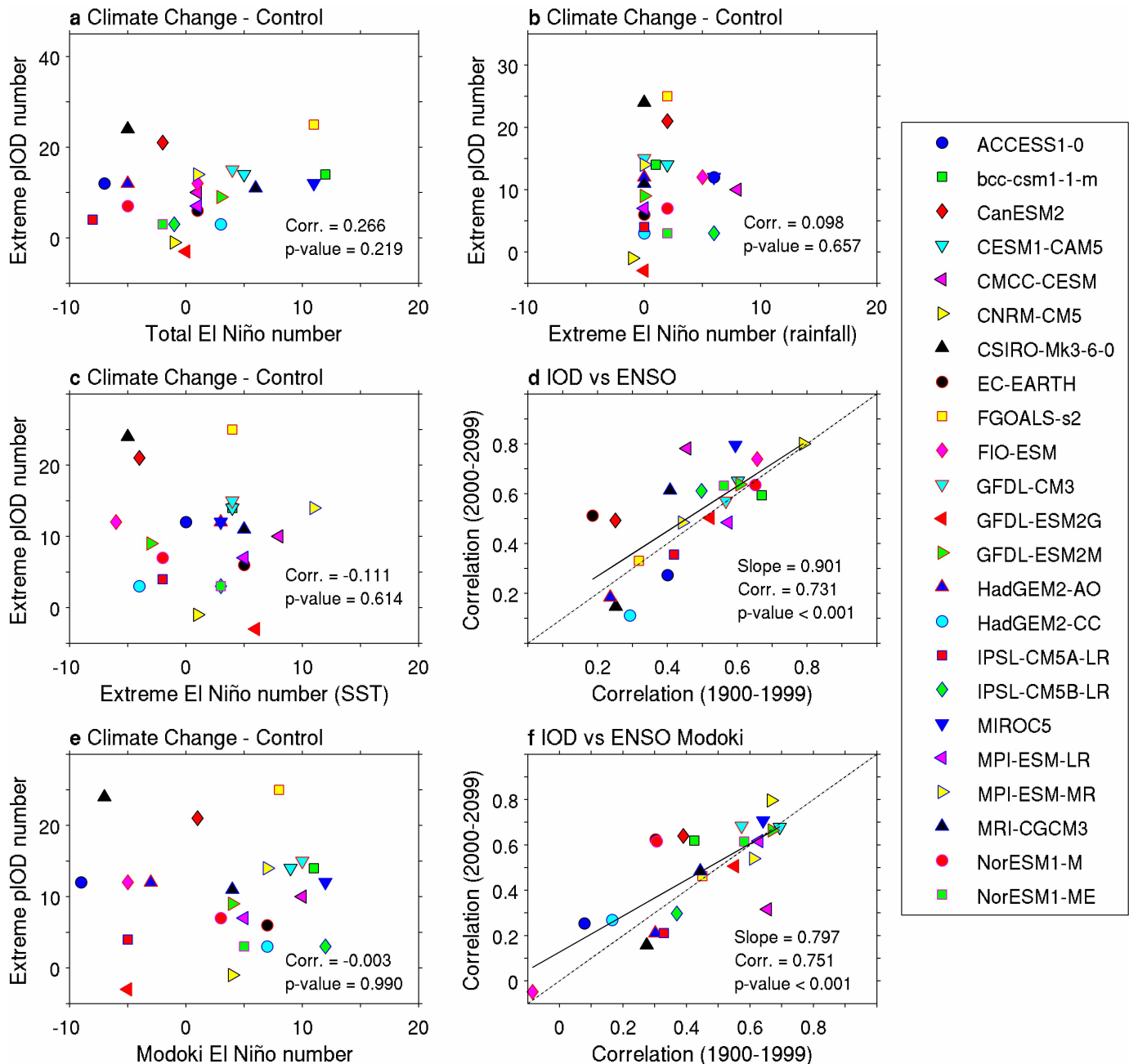
**Extended Data Figure 5 | Principal variability patterns of vertical velocity at 500 mb ( $\omega$ ), their nonlinear relationship, and the associated wind stress vectors during austral spring (September–November), based on a reanalysis<sup>33</sup>.** A positive vertical velocity indicates descending, while a negative  $\omega$  indicates ascending motion. **a, b**, Spatial patterns obtained by applying a statistical and signal processing method—EOF analysis—to the vertical velocity anomalies in the equatorial region ( $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ ,  $40^{\circ}\text{E}$ – $100^{\circ}\text{E}$ ) for data since 1979. The associated pattern and wind stress vectors from reanalysis data are obtained by linear regression onto the principal component time series of the EOFs. The first and second principal spatial pattern accounts for 32.6% and 16.8% of the total variance. The colour scale indicates vertical velocity in  $\text{Pa s}^{-1}$  per 1 s.d. change; blue or red contours indicate increased or decreased

convection. Note the different vector scales shown in the top right corner in **a** and **b**. **c**, A nonlinear relationship between the associated principal component time series. An extreme pIOD event (red dots) is defined as when the first principal component is greater than 1 s.d., and the second principal component is greater than 0.5 s.d. A moderate pIOD event (green dots) is determined from a detrended DMI when its amplitude is greater than 0.75 s.d., except for the 1994 and 1997 extreme pIOD events. Negative IOD and neutral years are indicated with blue dots. **d**, Relationship between the second principal component time series and rainfall over the eastern equatorial Pacific (Niño3) region ( $5^{\circ}\text{S}$ – $5^{\circ}\text{N}$ ,  $150^{\circ}\text{W}$ – $90^{\circ}\text{W}$ ). While the 1997 extreme pIOD was associated with a large rainfall in the Niño3 region, the 1961 and 1994 extreme pIODs were not.



**Extended Data Figure 6 | Multi-model ensemble average of the principal variability patterns of vertical velocity at 500 mb ( $\omega$ ), their nonlinear relationship, and the associated wind stress vectors during austral spring (September–November).** A positive vertical velocity indicates descending, while a negative  $\omega$  indicates ascending motion. **a, b,** Spatial patterns obtained by applying a statistical and signal processing method—EOF analysis—to the vertical velocity anomalies in the equatorial region ( $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ ,  $40^{\circ}\text{E}$ – $100^{\circ}\text{E}$ ). The associated pattern and wind stress vectors are obtained by linear regression onto the principal component time series. The colour scale below gives vertical velocity in  $\text{m s}^{-1}$  per 1 s.d. change; blue or red contours indicate increased or

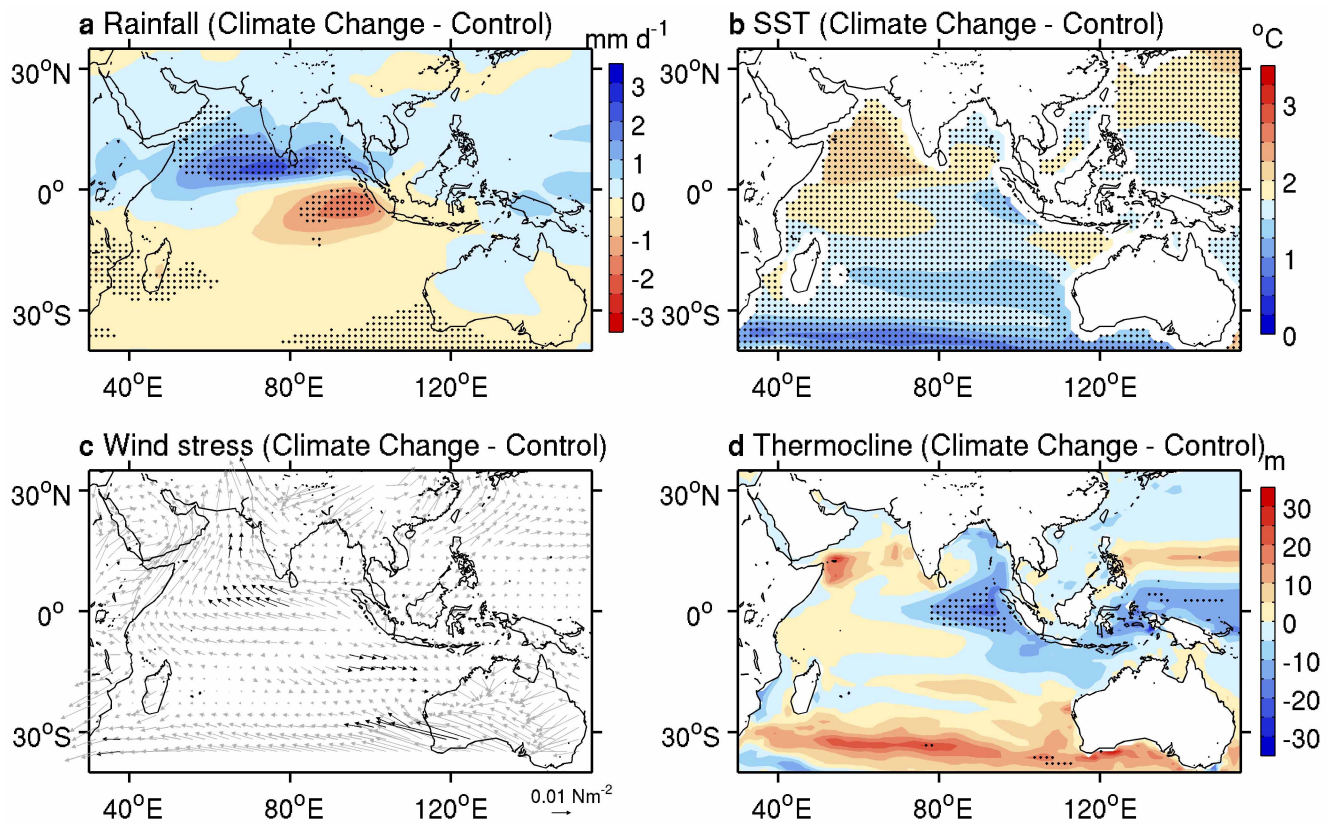
decreased convection. Note the different vector scales shown in the top right corner in **a** and **b**. **c** and **d,** A nonlinear relationship between the two principal component time series for the ‘control’ (1900–1999) and ‘climate change’ (2000–2099) periods. An extreme pIOD event (red dots) is defined as when the first principal component is greater than 1 s.d. and the second principal component is greater than 0.5 s.d. A moderate pIOD event (green dots) is determined from a detrended DMI when its amplitude is greater than 0.75 s.d. but is not an extreme pIOD event. Negative IOD and neutral years are indicated with blue dots. Number of extreme pIOD and moderate pIOD events is indicated in **c** and **d**.



**Extended Data Figure 7 | Multi-model statistics between El Niño and pIOD in selected CGCMs.** **a**, Changes ('climate change' minus 'control' period) in the number of occurrences of extreme pIOD events versus changes in the number of El Niño events defined as when the amplitude of the detrended Niño3 (5°S–5°N, 150°W–90°W) SST index is greater than 0.5 s.d. **b**, Changes in the number of extreme pIOD events versus changes in the number of El Niño events defined as when the Niño3 total rainfall is greater than 5 mm day<sup>-1</sup> as in ref. 38. **c**, The same as **b**, except an extreme El Niño is determined from a detrended Niño3 (5°S–5°N, 150°W–90°W) SST index when its amplitude is greater than 1.5 s.d. **d**, Correlation between a detrended Niño3 index and a detrended DMI index<sup>1</sup> for the 'climate change' (y axis) and the 'control' periods (x axis). **e**, Changes in the number of occurrences of extreme pIOD events

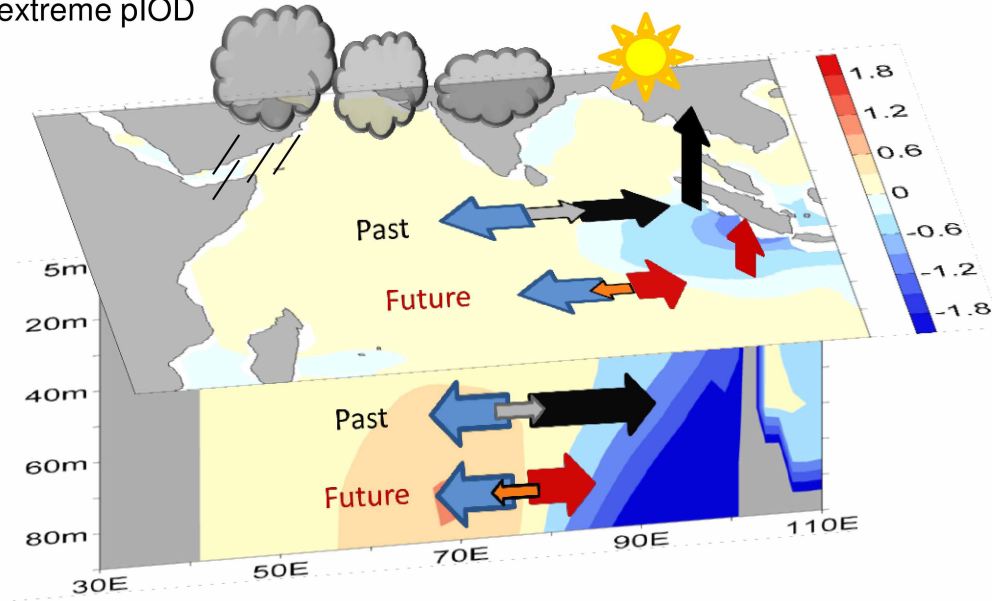
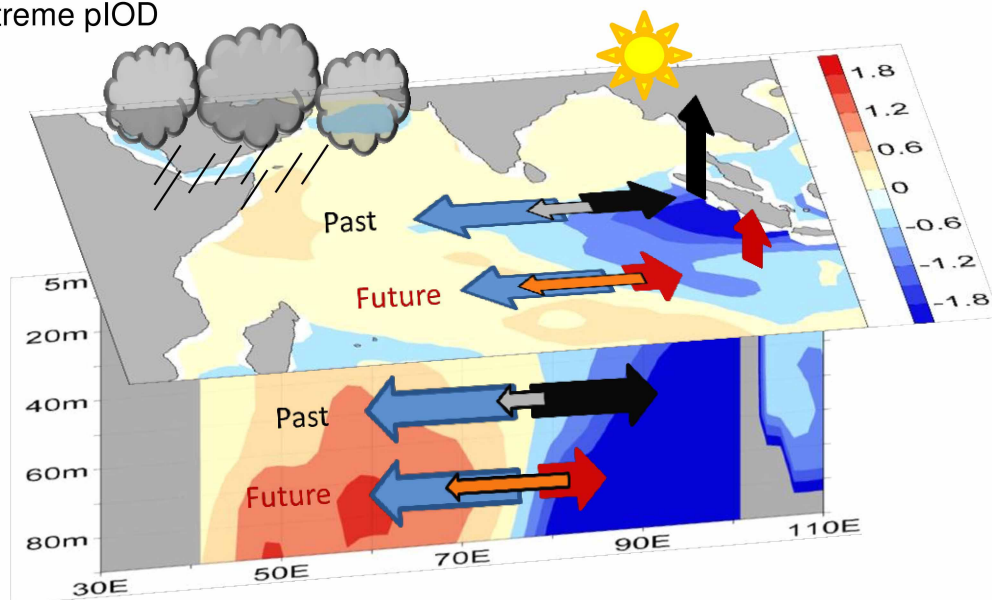
versus changes in the number of Modoki El Niño events defined as when the amplitude of a detrended index<sup>39</sup> (see Methods) is greater than 0.5 s.d. **f**, Correlation between a detrended El Niño index and a detrended DMI index for the 'climate change' (y axis) and the 'control' periods (x axis). The inter-model correlation and its statistical significance or otherwise are indicated in the bottom right corner of each panel, with a *P*-value less than 0.05, indicating significance above the 95% confidence level, a condition not met in **a**, **b**, **c** and **e**. Models with a stronger relationship between ENSO and the Indian Ocean dipole in the 'control' period tend to have a stronger such relationship in the 'climate change' period, and the tendency is statistically significant, although the relationship weakens slightly in the 'climate change' period. The same is true for the Modoki relationship between ENSO and the Indian Ocean dipole.





**Extended Data Figure 8 | Multi-model ensemble average of mean state changes for selected CGCMs.** The changes ('climate change' minus 'control' period) of the ensemble average mean state for: **a**, rainfall ( $\text{mm day}^{-1}$ ), **b**, SST ( $^{\circ}\text{C}$ ), **c**, wind stress vectors ( $\text{N m}^{-2}$ ) and **d**, thermocline depth (m). The result shows that rainfall off Sumatra is decreasing, the southern eastern Indian Ocean

is warming at a slower rate than the west, there is a trend of easterlies over the equatorial Indian Ocean, and the thermocline is shallowing in the eastern equatorial Indian Ocean. Areas where changes are statistically significant at the 95% confidence level are indicated with stipples, in **a**, **b**, and **d**. In **c**, vectors in bold indicate statistical significance at the 95% confidence level.

**a** Non-extreme pIOD**b** Extreme pIOD

**Extended Data Figure 9 | Schematic of extreme pIOD in response to greenhouse warming.** **a**, pIOD events are characterized by westward-flowing wind anomalies (blue arrow at the surface) and the associated westward-flowing current anomalies (blue arrow at depth) acting against the prevailing background eastward circulations (black arrows), in association with the anomalous positive west-minus-east SST gradient. These result in generally weaker-than-normal eastward atmosphere and ocean circulations (grey

arrows), with anomalously wet condition in the west and dry in the east. **b**, During extreme pIOD events, these anomalies are amplified, with occurrences of strong reversals of the mean eastward winds and currents (grey arrows). As the mean Walker circulation and the associated eastward-flowing ocean current weaken (red arrows), wind and current reversals (orange arrows) can occur more easily in association with pIOD anomalies. Greenhouse warming thus induces more frequent occurrences of extreme pIOD events.

Extended Data Table 1 | Performance of 23 selected CMIP5 CGCMs forced under climate change emission scenario RCP8.5

CGCM	IODE skewness 1900-1999	Extreme pIOD events defined as (EOF1>1 s.d. & EOF2>0.5 s.d.) (1900-1999)/(2000-2099)	Extreme pIOD events defined as (EOF1>1.5 s.d. & EOF2>0.5 s.d.) (1900-1999)/(2000-2099)
ACCESS1-0 SST $\omega$ $\tau$	-0.2606	5/17	4/10
bcc-csm1-1-m SST $\omega$	-0.7997	4/18	1/17
CanESM2 SST $\omega$ $\tau$	-1.4207	1/22	1/22
CESM1-CAM5 SST $\omega$	-0.5826	8/22	8/22
CMCC-CESM SST $\omega$	-1.2214	3/13	2/10
CNRM-CM5 SST $\omega$ $\tau$	-0.2169	13/12	5/8
CSIRO-Mk3-6-0 SST $\omega$ $\tau$	-0.3824	2/26	2/26
EC-EARTH SST	-0.6927	6/12	6/12
FGOALS-s2 SST $\omega$ $\tau$	-0.3378	4/29	2/25
FIO-ESM SST $\omega$	-1.6906	4/16	4/14
GFDL-CM3 SST $\omega$ $\tau$	-0.8189	5/20	5/20
GFDL-ESM2G SST $\omega$	-1.1149	8/5	6/4
GFDL-ESM2M SST $\omega$ $\tau$	-0.9843	9/18	8/17
HadGEM2-AO SST $\omega$	-0.2170	5/17	3/16
HadGEM2-CC SST $\tau$	-0.2343	10/13	8/12
IPSL-CM5A-LR SST $\omega$ $\tau$	-0.7191	6/10	4/8
IPSL-CM5B-LR SST $\omega$	-0.1511	9/12	3/11
MIROC5 SST $\omega$ $\tau$	-1.7136	0/12	0/11
MPI-ESM-LR SST $\omega$	-2.6676	5/12	4/12
MPI-ESM-MR SST $\omega$	-1.3712	4/18	4/18
MRI-CGCM3 SST $\omega$ $\tau$	-0.6987	5/16	4/14
NorESM1-M SST $\omega$ $\tau$	-0.5208	7/14	6/10
NorESM1-ME SST $\omega$	-0.5174	10/13	4/5
<b>Total and % change between the two periods (23models)</b>		<b>133/367(176%) 17.3/6.3 Years</b>	<b>94/324(245%) 24.5/7.1 Years</b>

These CGCMs are selected in terms of SST skewness in the eastern pole of the Indian Ocean dipole (IODE) and each model's ability to simulate the nonlinear relationship between rainfall EOF1 and EOF2 (Fig. 2). The sensitivity of changes in extreme pIOD events from the 'control' period to the 'climate change' period to different definitions is tested. An extreme pIOD event is defined as when the first principal component time series is greater than 1 s.d., or 1.5 s.d., and the second principal component time series is greater than 0.5 s.d. Numbers in red type indicate a decrease from the 'control' period (1900–1999) to 'climate change' period (2000–2099). Multi-model average SST skewness in the eastern pole of the Indian Ocean dipole is  $-0.84$ , compared with the observed value of  $-0.85$ . The subscripts SST,  $\omega$  and  $\tau$  indicate that the data of SST, vertical velocity at 500 mb and surface wind stress are available, respectively.



# Elevated CO<sub>2</sub> further lengthens growing season under warming conditions

Melissa Reyes-Fox<sup>1\*</sup>, Heidi Steltzer<sup>2\*</sup>, M. J. Trlica<sup>3</sup>, Gregory S. McMaster<sup>4</sup>, Allan A. Andales<sup>5</sup>, Dan R. LeCain<sup>6</sup> & Jack A. Morgan<sup>6</sup>

Observations of a longer growing season through earlier plant growth in temperate to polar regions have been thought to be a response to climate warming<sup>1–5</sup>. However, data from experimental warming studies indicate that many species that initiate leaf growth and flowering earlier also reach seed maturation and senesce earlier, shortening their active and reproductive periods<sup>6–10</sup>. A conceptual model to explain this apparent contradiction<sup>11</sup>, and an analysis of the effect of elevated CO<sub>2</sub>—which can delay annual life cycle events<sup>12–14</sup>—on changing season length, have not been tested. Here we show that experimental warming in a temperate grassland led to a longer growing season through earlier leaf emergence by the first species to leaf, often a grass, and constant or delayed senescence by other species that were the last to senesce, supporting the conceptual model. Elevated CO<sub>2</sub> further extended growing, but not reproductive, season length in the warmed grassland by conserving water, which enabled most species to remain active longer. Our results suggest that a longer growing season, especially in years or biomes where water is a limiting factor, is not due to warming alone, but also to higher atmospheric CO<sub>2</sub> concentrations that extend the active period of plant annual life cycles.

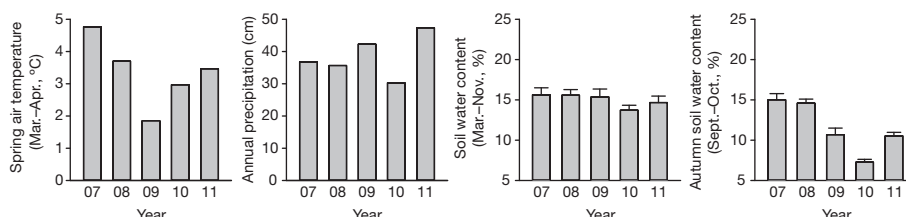
Climate varied considerably between the years studied in this experiment, most notably through a cool spring in 2009, greater precipitation than usual in 2009 and 2011, and low autumn soil water content in 2010 (Fig. 1 and Extended Data Figs 1, 2). Nonetheless, in all but one year, climate warming conditions (*cT*, where *c* is a condition of relatively low CO<sub>2</sub> and *T* is relatively high temperature) changed the timing of species' annual life cycles, increasing the length of the growing season and the reproductive season relative to control conditions (*ct*, where *t* is relatively low temperature) (Fig. 2). Warming led to earlier timing of events by the first species to leaf or flower in most years, but species' sensitivity to warming varied among species and among years (Fig. 3 and Extended Data Table 1). Often, a cool-season grass, *Koeleria macrantha*, was the first species to leaf and the first species to flower in control (*ct*) and warmed (*cT*) plots. In most years, warming advanced leaf emergence and flowering of *K. macrantha* (Fig. 3; see Supplementary Information for timing of annual life cycle events for all species in all treatments in all years), yet, in 2009, a year characterized by a cool spring (Fig. 1), warming delayed leaf

emergence of *K. macrantha* by 9 days. Warming delays leafing and flowering for some species<sup>7,15</sup>, although the mechanism behind this is not clear.

Contrasting species' responses to warming between years limits interpretation of 5-year means; thus we present data yearly. Additionally, yearly data illustrate that the first and last species to complete annual life cycle events that determine the start and end of the growing or reproductive seasons shifted between treatments and between years (Fig. 3). These shifts indicate complementarity among species in response to interannual climate variation and warming, countering the tendency of warming to shorten the growing season. For example, in 2009, when warming (*cT*) delayed leaf emergence of *K. macrantha*, leaf emergence of *Artemisia frigida*, a sub-shrub, was not affected, leading to a shift in which species was the first to leaf and no change in timing for the start of the growing season relative to the control (*ct*) (Fig. 3).

Warming (*cT*) led to earlier leaf emergence and flowering, but also to earlier seed maturation and canopy senescence for some species, especially *K. macrantha*, relative to the control (*ct*) (Fig. 3). Seed maturation by *A. frigida*, consistently the last species to complete this event, was not affected by warming. Therefore, a longer reproductive season in response to warming primarily resulted from earlier flowering by *K. macrantha*. The mean active period for *K. macrantha* shortened (Fig. 4 and Extended Data Table 2), but longer growing seasons resulted, because *A. frigida* and *Hesperostipa comata* did not change or delayed the timing of canopy senescence (Fig. 3). Warming extended the duration of the mean reproductive period over the 5 years for three of the six species, including *K. macrantha* (Fig. 4 and Extended Data Table 2), primarily through lengthening the reproductive period in 2011, the year with the most precipitation (Fig. 1 and Supplementary Information). Delayed canopy senescence due to warming and a later end of the growing season also only occurred in 2011 (Figs 2 and 3).

Variation in plant life history traits within the grassland, such as early season growth versus late season tissue maintenance, led to differences in species' responses to warming, supporting a conceptual model of how individual species' responses determine growing season length<sup>11</sup>. However, in contrast to the model's prediction, divergence in species' active periods was small. Several species, including *H. comata*, lengthened their

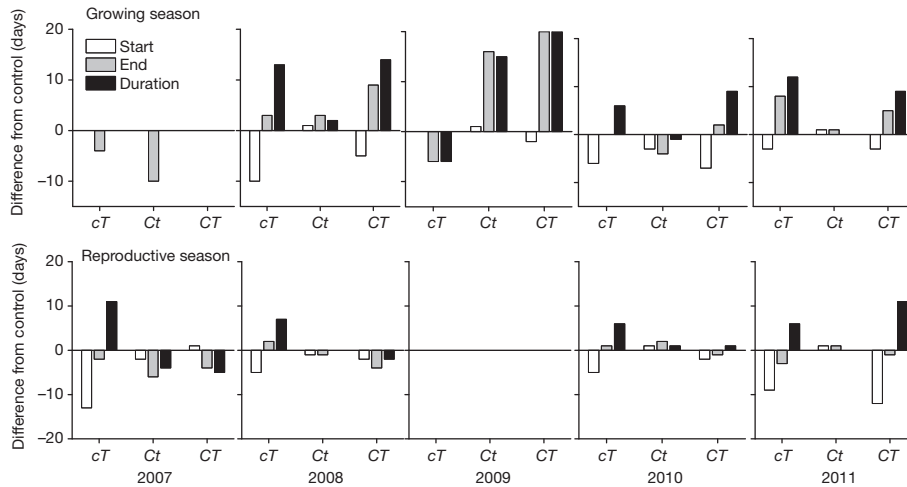


**Figure 1 | Interannual variation in climate and microclimate (2007–2011).** Values are mean spring air temperature for day of year (DOY) 60–120 and annual precipitation for the study site, and mean growing season (DOY 60–334) and autumn (DOY 244–304) soil water content for the control plots

(means  $\pm$  1 standard error of the mean (s.e.m.),  $n = 5$  plots). Spring air temperature and autumn soil water content correspond with timing of leaf emergence and canopy senescence, respectively.

<sup>1</sup>USDA-ARS, Soil Plant Nutrient Research Unit and Northern Plains Area, Fort Collins, Colorado 80526, USA. <sup>2</sup>Department of Biology, Fort Lewis College, Durango, Colorado 81301, USA. <sup>3</sup>Department of Forest and Rangeland Stewardship, Colorado State University, Fort Collins, Colorado 80523, USA. <sup>4</sup>USDA-ARS, Agricultural Systems Research Unit and Northern Plains Area, Fort Collins, Colorado 80526, USA. <sup>5</sup>Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado 80523, USA. <sup>6</sup>USDA-ARS, Rangeland Resources Research Unit, Fort Collins, Colorado 80526, USA.

\*These authors contributed equally to this work.

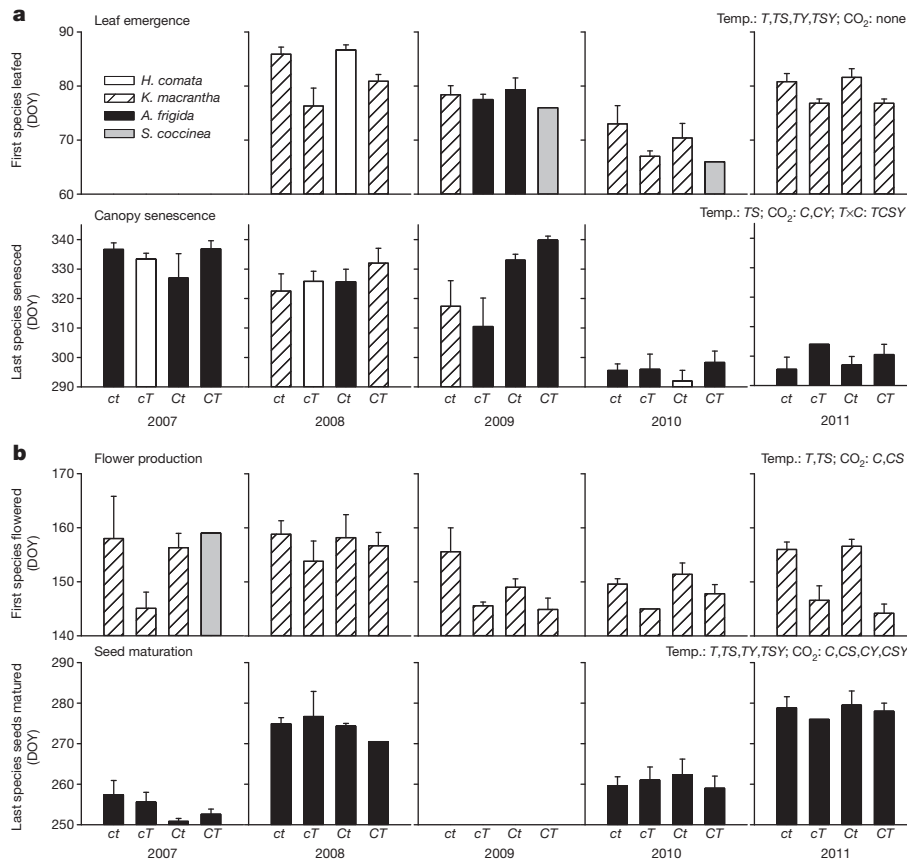


**Figure 2 | Effect of warming and elevated CO<sub>2</sub> on growing and reproductive season length (2007–2011).** Values are the mean number of days difference between treatment and control conditions for start, end and duration; see Fig. 3

active period in response to warming (Fig. 4), maintaining continuity of seasonal growth by the plant community. Similarly, for several species reproductive periods lengthened under warming (Fig. 4), limiting divergence within the reproductive season. Thus, our results contrast with other studies in which experimental warming led to divergent flowering responses between species<sup>9</sup> and mid-season, low floral abundance as the climate warmed<sup>16</sup>. Longer active and reproductive periods by at least some species would reduce the adverse effects of warming on trophic interactions and ecosystem function<sup>17,18</sup>.

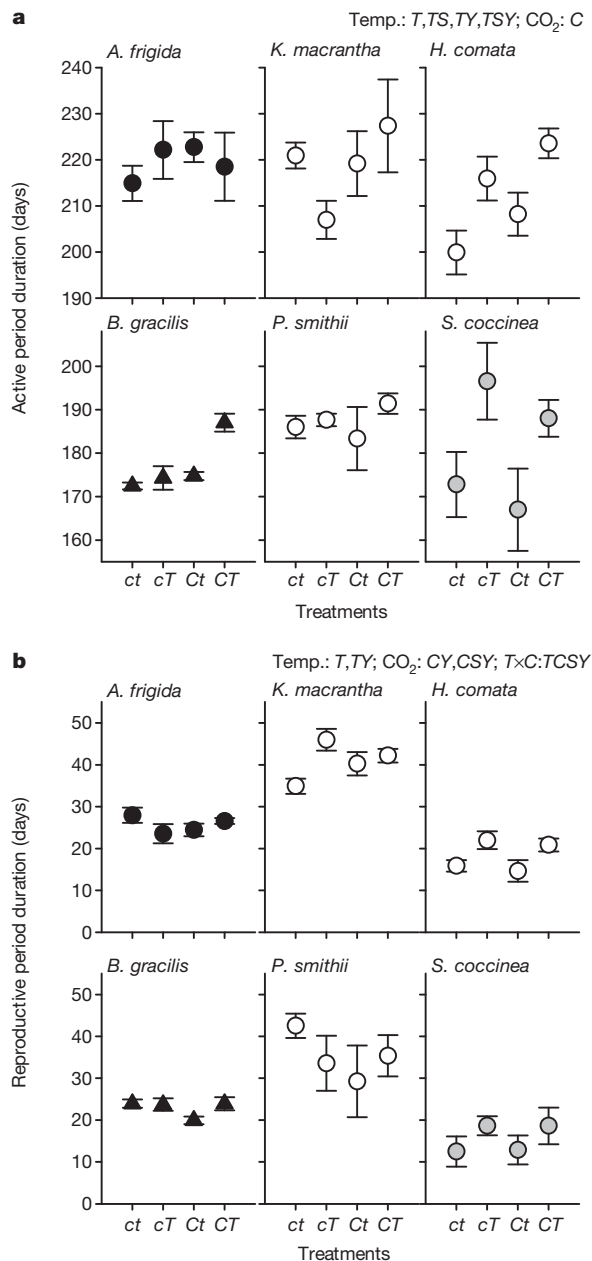
for corresponding s.e.m. and significant effects ( $n = 5$  plots). Negative values indicate earlier onset of events or shortening of growing or reproductive season length. See Methods for explanation of missing data.

In conditions of elevated CO<sub>2</sub> (CT, where C is relatively high CO<sub>2</sub>), growing season duration was further lengthened relative to warming alone (cT) through the delay of canopy senescence (Figs 2, 3 and Extended Data Table 1). In 2009, when spring was cool and annual precipitation was high, elevated CO<sub>2</sub> extended the growing season by delaying senescence of *A. frigida* by 29 days in the warmed ecosystem. Although the magnitude of the response was less, the growing season was also significantly increased in 2008, when under conditions of elevated CO<sub>2</sub> senescence occurred 6 days later in the warmed ecosystem. On average during our

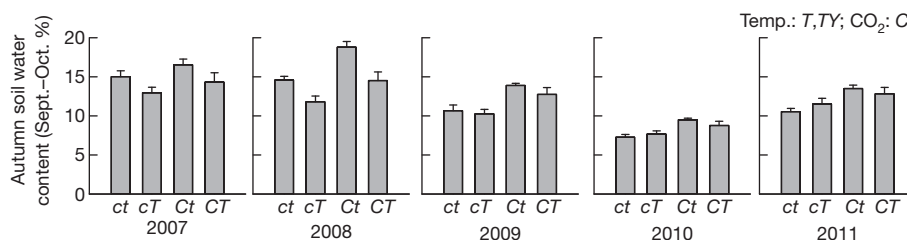


**Figure 3 | Effect of warming and elevated CO<sub>2</sub> on timing of annual life cycle events (2007–2011).** a, b, Values are day of year (DOY) for start and end of growing (a) and reproductive seasons (b) (means  $\pm$  1 s.e.m.,  $n = 5$  plots). Significant effects ( $P < 0.05$ , two-sided) from a four-way ANOVA testing

temperature (Temp.; T), elevated CO<sub>2</sub> (C), species (S), year (Y) and all interactions are reported at the top of each row of panels. Species, year and their interactions were highly significant for all events ( $P < 0.001$ ) and are not listed. See Extended Data Table 1 for complete ANOVA results.



**Figure 4 | Effect of warming and elevated CO<sub>2</sub> on the duration of species' active and reproductive periods.** a, b, Values are means across years (2007–2011) for species' active period duration (a) and reproductive period duration (b) (means  $\pm$  1 s.e.m.,  $n$  = 5 plots). Sub-shrub is indicated by filled black circles, cool-season grasses by open circles, warm-season grass by filled triangles, forb by filled grey circles. Significant effects ( $P$  < 0.05, two-sided) are reported as in Fig. 3 at the top of each group of panels. See Extended Data Table 2 for complete ANOVA results.



**Figure 5 | Effect of warming and elevated CO<sub>2</sub> on autumn soil water content (5–25 cm, September–October, 2007–2011).** Values are means for soil water content corresponding with the timing of canopy senescence (means  $\pm$  1 s.e.m.,

5-year study, the growing season ended 7.6 days later due to warming and elevated CO<sub>2</sub> (CT) relative to warming alone (cT), and was 14.2 days longer. Owing to warming alone (cT versus ct), the growing season began 4.7 days earlier and was 6.2 days longer on average. The effect of warming alone probably cannot account for the observed change in growing season length of  $\sim$ 2 to 5 days per decade during the mid- to late twentieth century<sup>1–5</sup>.

Our results demonstrate that the effects of warming and elevated CO<sub>2</sub> (CT) on annual life cycle events that determine growing and reproductive season lengths depend on climate, varying in the magnitude and even direction of response between years (Figs 2, 3 and Extended Data Table 1). For example, in 2010, a year with low late summer and autumn precipitation and low autumn soil water content (Fig. 1), elevated CO<sub>2</sub> (Ct and CT) did not affect canopy senescence (Fig. 3). Although elevated CO<sub>2</sub> (Ct and CT) led to greater autumn soil water content in all years (Fig. 5 and Extended Data Table 3), in 2010 the water savings may not have been sufficient to increase water availability above a threshold ( $\sim$ 10%) that corresponds to the permanent wilting point (11%) (ref. 19). In the warmed ecosystem, elevated CO<sub>2</sub> (CT) led to the greatest increase in autumn soil water content in 2008 and 2009 (Fig. 5), the years in which the greatest delays in canopy senescence due to elevated CO<sub>2</sub> (CT) occurred (Fig. 3).

Increasing the CO<sub>2</sub> concentration in the warmed ecosystem (CT versus cT) lengthened the active period of all grass species (Fig. 4 and Extended Data Table 2). Even the early growing grass *K. macrantha*, which senesced early due to warming alone (cT), showed delayed senescence under elevated CO<sub>2</sub> and warming conditions (CT). Warming and elevated CO<sub>2</sub> (CT) led to a shorter reproductive season in 3 of 4 years relative to warming alone (cT) by decreasing the advance in flowering date and earlier seed maturation by *A. frigida* in 2007 and 2008 (Figs 2 and 3). Elevated CO<sub>2</sub> tended to decrease or have no effect on species' reproductive period (Fig. 4 and Extended Data Table 2). Thus, in the warmed ecosystem, elevated CO<sub>2</sub> (CT versus cT) caused species to remain active longer after seed maturation, which would not benefit fitness in the year in question but may affect it in consequent years.

Altered flowering times and species' reproductive periods may have long-term consequences for plants and other trophic levels. Other studies have also found that elevated CO<sub>2</sub> has a greater effect on flowering times under warming conditions, as well as a greater effect on late-flowering species<sup>14,20</sup>. Our data indicate that higher atmospheric CO<sub>2</sub> concentrations may be contributing to observed changes over time in flowering patterns, such as a shorter reproductive season and greater asynchrony with pollinators, which have previously been attributed to warming<sup>16,17</sup>.

Furthermore, the dominant hypothesis among global change ecologists is that Earth's longer growing seasons are due to climate warming alone. Our results suggest that this hypothesis needs modification to incorporate the effects of elevated CO<sub>2</sub>. We provide evidence in multiple years and of a mean effect over the 5-year study that elevated CO<sub>2</sub> further increases growing season length in a warmed, temperate plant community. In many ecosystems, sufficient water availability is needed to sustain plant tissues from summer into autumn. Dry conditions during warm years have led to early senescence and even the death of long-lived plants<sup>21,22</sup>. Elevated CO<sub>2</sub> counteracts the negative effect of warming on water availability<sup>19,23</sup> (Fig. 5), often delaying the timing of plant life cycle events<sup>12,13</sup> (Fig. 3). The effects of

$n$  = 5 plots). Significant effects ( $P$  < 0.05, two-sided) are reported as in Fig. 3. See Extended Data Table 3 for complete ANOVA results.

warming and elevated CO<sub>2</sub> vary across species, events and years. At the community level, the different responses lead to a longer growing season in most years, because only one species within the plant community needs to leaf earlier for spring to begin earlier and a different species can senesce later to extend autumn<sup>11</sup>. We demonstrate that warming fairly consistently leads to earlier growth in spring and elevated CO<sub>2</sub> to later senescence in autumn, with both mechanisms leading to a longer growing season.

The stature of grasslands and their ability to encompass thousands of individual plants, many species, and different growth forms within a relatively small area make them ideal ecosystems in which to conduct global change experiments<sup>9,12,23</sup>. For example, in our experiment, the varied responses of the three cool-season grasses, *K. macrantha*, *H. comata* and *Pascopyrum smithii*, a species that did not affect growing or reproductive season length in any year, suggest that different species, even within a growth form, respond in unique ways to warming and elevated CO<sub>2</sub> (Figs 3 and 4). As a result, plant community response did not depend on a specific species or growth form, and we expect that responses would be similar in other temperate to polar plant communities, especially in years or biomes where water is a limiting factor.

Although considerably less than the ~200 p.p.m. of CO<sub>2</sub> enrichment that occurred in our experiment, the ~60 p.p.m. increase in global atmospheric CO<sub>2</sub> concentrations since the 1970s is probably sufficient to elicit significant stomatal closure<sup>24,25</sup>, resulting in some water savings and an effect on phenology, as in our experiment. Certainly the ~115 p.p.m. increase in global atmospheric CO<sub>2</sub> concentrations since industrialization has been more than enough to elicit considerable CO<sub>2</sub>-induced water savings<sup>23</sup> and affect growing season length, although data on growing season length are available primarily from the mid-twentieth century<sup>1–5</sup>. Ongoing increases in ambient CO<sub>2</sub> are expected to continue to shift the timing of species' reproductive periods and senescence, and thus the duration of the growing and reproductive seasons.

## METHODS SUMMARY

The experiment, initiated in 2006, is located in temperate grassland in Wyoming, United States (41° 11' N, 104° 54' W). It includes two levels of temperature (ambient and warmed, 1.5 or 3.0 °C warmer during the day and night, respectively) and two levels of atmospheric CO<sub>2</sub> concentrations (ambient and elevated, 385 p.p.m.v. and 600 p.p.m.v. CO<sub>2</sub>, respectively) in a factorial combination with five replicate plots per treatment. T-FACE technology was used for increasing the temperature<sup>26</sup>. Free air CO<sub>2</sub> enrichment (FACE) technology was used for elevating CO<sub>2</sub> (ref. 27). Further description of the experiment can be found elsewhere<sup>19</sup>.

The timing of the four annual life cycle events that determine the start and end of species' active and reproductive periods (leaf emergence, flower production, seed maturation and canopy senescence) was observed weekly for six common species. The most abundant species in each growth form were chosen. From 2007 to 2009, individual plants were marked upon emergence and monitored for the duration of the growing season. The timing of each event was characterized by the mean value for the marked individuals of each species. In 2010 and 2011, the timing of each event was characterized by the median value, the point at which an event was completed by half the typical number of marked individuals for each species within a plot. Further description is included in the Methods.

We analysed the data across the years through an analysis of variance (ANOVA) for each life cycle event, duration of active and reproductive periods, and autumn soil water content, using Proc Mixed to test the main effects of temperature, elevated CO<sub>2</sub>, species (if applicable) and year, and all their interactions (SAS version 9.2). Soil type was included as a random effect and was not significant. Mixed model ANOVA allows for unequal variances and data were near a normal distribution. Analyses were done on untransformed data.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 November 2013; accepted 5 March 2014.

Published online 23 April 2014.

- Kim, Y., Kimball, J. S., Zhang, K. & McDonald, K. C. Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: implications for regional vegetation growth. *Remote Sens. Environ.* **121**, 472–487 (2012).

- Menzel, A. *et al.* European phenological response to climate change matches the warming pattern. *Glob. Change Biol.* **12**, 1969–1976 (2006).
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. Increased plant growth in the northern latitudes from 1981 to 1991. *Nature* **386**, 698–702 (1997).
- Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003).
- Schwartz, M. D., Ahas, R. & Aasa, A. Onset of spring starting earlier across the Northern Hemisphere. *Glob. Change Biol.* **12**, 343–351 (2006).
- Hoffmann, A. A. *et al.* Phenological changes in six Australian subalpine plants in response to experimental warming and year-to-year variation. *J. Ecol.* **98**, 927–937 (2010).
- Hollister, R. D., Webber, P. J. & Bay, C. Plant response to temperature in northern Alaska: implications for predicting vegetation change. *Ecology* **86**, 1562–1570 (2005).
- Post, E. S., Pedersen, C., Wilms, C. C. & Forchhammer, M. C. Phenological sequences reveal aggregate life history response to climatic warming. *Ecology* **89**, 363–370 (2008).
- Sherry, R. A. *et al.* Divergence of reproductive phenology under climate warming. *Proc. Natl Acad. Sci. USA* **104**, 198–202 (2007).
- Zavaleta, E. S. *et al.* Plants reverse warming effect on ecosystem water balance. *Proc. Natl Acad. Sci. USA* **100**, 9892–9893 (2003).
- Steltzer, H. & Post, E. Seasons and life cycles. *Science* **324**, 886–887 (2009).
- Cleland, E. E., Chiariello, N. R., Loarie, S. R., Mooney, H. A. & Field, C. B. Diverse responses of phenology to global changes in a grassland ecosystem. *Proc. Natl Acad. Sci. USA* **103**, 13740–13744 (2006).
- Körner, C. Plant CO<sub>2</sub> responses: an issue of definition, time and resource supply. *New Phytol.* **172**, 393–411 (2006).
- Springer, C. J. & Ward, J. K. Flowering time and elevated atmospheric CO<sub>2</sub>. *New Phytol.* **176**, 243–255 (2007).
- Fitter, A. H. & Fitter, R. S. R. Rapid changes in flowering time in British plants. *Science* **296**, 1689–1691 (2002).
- Aldridge, G., Inouye, D. W., Forrest, J. R. K., Barr, W. A. & Miller-Rushing, A. J. Emergence of a mid-season period of low floral resources in a montane meadow ecosystem associated with climate change. *J. Ecol.* **99**, 905–913 (2011).
- Høye, T. T., Post, E., Schmidt, N. M., Trøjsgaard, K. & Forchhammer, M. C. Shorter flowering seasons and declining abundance of flower visitors in a warmer Arctic. *Nature Climate Change* **3**, 759–763 (2013).
- Richardson, A. D. *et al.* Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agric. For. Meteorol.* **169**, 156–173 (2013).
- Morgan, J. A. *et al.* C<sub>4</sub> grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. *Nature* **476**, 202–205 (2011).
- Hovenden, M. J., Wills, K. E., Vander Schoor, J. K., Williams, A. L. & Newton, P. C. D. Flowering phenology in a species-rich temperate grassland is sensitive to warming but not elevated CO<sub>2</sub>. *New Phytol.* **178**, 815–822 (2008).
- Breshears, D. D. *et al.* Regional vegetation die-off in response to global-change-type drought. *Proc. Natl Acad. Sci. USA* **102**, 15144–15148 (2005).
- van Mantgem, P. J. *et al.* Widespread increase of tree mortality rates in the western United States. *Science* **323**, 521–524 (2009).
- Fay, P. A. *et al.* Soil-mediated effects of subambient to increased carbon dioxide on grassland productivity. *Nature Climate Change* **2**, 742–746 (2012).
- Ainsworth, E. A. & Rogers, A. The response of photosynthesis and stomatal conductance to rising [CO<sub>2</sub>]: mechanisms and environmental interactions. *Plant Cell Environ.* **30**, 258–270 (2007).
- Ball, J. T., Woodrow, I. E. & Berry, J. A. In *Progress in Photosynthesis Research* (ed. Biggins, J.) 221–224 (Martinus-Nijhoff, 1987).
- Kimball, B. A. *et al.* Infrared heater arrays for warming ecosystem field plots. *Glob. Change Biol.* **14**, 309–320 (2008).
- Miglietta, F. *et al.* Free-air CO<sub>2</sub> enrichment (FACE) of a poplar plantation: the POPFACE fumigation system. *New Phytol.* **150**, 465–476 (2001).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The following individuals contributed to the installation and maintenance of the Prairie Heating and CO<sub>2</sub> Enrichment (PHACE) project: D. Smith, D. Blumenthal, E. Pendall, E. Hardy, L. Griffith, A. Hansen, K. Corp, V. Banuelos, G. Tinnin, M. West, C. Brooks, M. Busick, D. Milchunas, G. Dunn and L. Ahuja. Funding for this work was supported by the US Department of Agriculture Agricultural Research Center (USDA-ARS) Climate Change, Soils & Emissions Program, by the US Department of Energy's Office of Science through the Terrestrial Ecosystem Science Program, by the National Science Foundation (DEB no. 1021559) and by Colorado State University. D. Inouye provided comments that improved the manuscript.

**Author Contributions** M.R.-F., M.J.T., A.A.A., G.S.M. and J.A.M. designed the research. M.R.-F. and D.R.L. conducted the observations. J.A.M. oversaw the PHACE experiment. H.S. and M.R.-F. analysed the data and wrote the manuscript. All authors contributed to revision of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.S. ([steltzer\\_h@fortlewis.edu](mailto:steltzer_h@fortlewis.edu)) or M.R.-F. ([Melissa.reyes@ars.usda.gov](mailto:Melissa.reyes@ars.usda.gov)).



## METHODS

**Site description.** The Prairie Heating and Carbon Dioxide Enrichment (PHACE) experiment, initiated in 2006, is located west of Cheyenne, Wyoming, United States at the USDA-ARS High Plains Grasslands Research Station in the US Great Plains (41° 11' N, 104° 54' W, elevation 1,930 m). This is a Northern mixed-grass prairie ecosystem, with a plant community composed of 55% cool-season grasses, 25% warm-season grasses, and 20% sedges, forbs and small shrubs. Total annual precipitation averages 38.5 cm and mean daily air temperatures range from  $-2.5^{\circ}\text{C}$  in January to  $17.5^{\circ}\text{C}$  in July. The average wind speed is  $6\text{ m s}^{-1}$  with gusts up to  $35\text{ m s}^{-1}$ . The site comprises two distinct soil types: an Ascalon variant loam (fine loamy, mixed mesic) at the north end of the field and an Altvan loam (fine loamy over sandy, mixed mesic) on the south end. The site has a history of moderate grazing from 1928 until the PHACE project began in 2006.

**Experimental design.** The experiment includes two levels of temperature (ambient and warmed, 1.5 and  $3.0^{\circ}\text{C}$  warmer during the day and night, treatments *t* and *T*, respectively) and two levels of atmospheric  $\text{CO}_2$  concentrations (ambient 385 p.p.m.v. and elevated 600 p.p.m.v.  $\text{CO}_2$ , treatments *c* and *C*, respectively) in a factorial combination with five replicate plots per treatment (*ct*, *cT*, *Ct* and *CT*) for a total of 20 plots. Warming and elevated  $\text{CO}_2$  treatments were randomly assigned to the 3.3 m diameter circular plots. T-FACE technology for increasing temperature was implemented on 10 April 2007, after leaf emergence by cool-season grasses and shrubs, and warmed plots year round for the duration of the experiment<sup>26</sup>. As warming began after the first species leafed in 2007, leaf emergence data were omitted for all species in 2007 and growing season length was not calculated. Dummy heaters were installed in non-heated plots to eliminate response differences that may result from shading or other influences caused by the heating apparatuses. Free air  $\text{CO}_2$  enrichment (FACE) technology was used for elevating  $\text{CO}_2$  and began in 2006 (ref. 27). The  $\text{CO}_2$  fumigation system ran continuously during the growing season. The warming treatment effectively accelerated the accumulation of growing degree days in all years (Extended Data Fig. 1). When placed in the historical context of the last century, below average precipitation fell in 2007, 2008 and 2010 with above average precipitation falling in 2009 and 2011. Further description of the experiment, including the instrumentation used for monitoring climate and microclimate, is available elsewhere<sup>19</sup>.

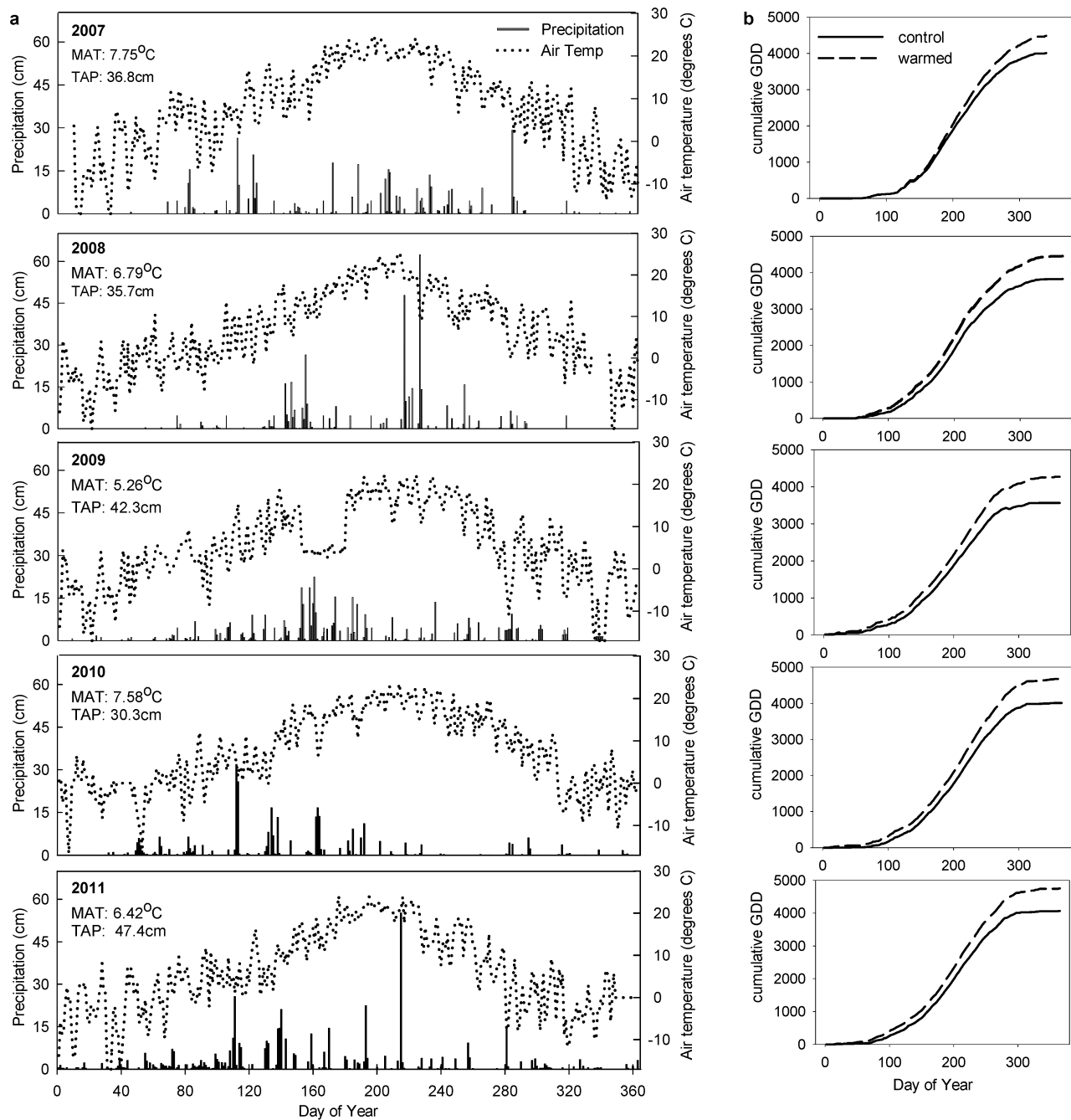
**Phenology observations.** The timing of four life cycle events that determine the start and end of species' active and reproductive periods (leaf emergence, flower production, seed maturation and canopy senescence) was observed weekly for six common species. The most abundant species in each growth form were chosen, including the one sub-shrub (*A. frigida*, L.), a warm-season grass (*Bouteloua gracilis*, Lag. ex Griffiths), three cool-season grasses (*H. comata* (Elias) Barkworth; *K. macrantha* (Ledeb.) Schult; and *P. smithii* (Rydb.) Á. Löve), and a widespread forb (*Sphaeralcea coccinea* (Nutt.) Rydb.). Leaf emergence was characterized by the first new, green leaf to appear on a shoot. The first open flower (forb and sub-shrub) or inflorescence to emerge from the leaf sheath

(grasses) determined the timing of flower production. Flower desiccation (forb and sub-shrub) and seed head colour and release (grasses) identified the timing of seed maturation. And the timing of canopy senescence was characterized by the first sign of full canopy leaf death or loss.

From 2007 to 2009, individual plants of each species were marked upon emergence (typically 12) and monitored for the duration of the growing season. The timing of an event was characterized by the mean value (day of year (DOY)) for the marked individuals of each species. Some species did not flower in all years. In 2009, incomplete data were collected on reproductive life cycle events, so reproductive season length was not determined. In 2010 and 2011, the timing of an event was characterized as the point at which a minimum number of individuals (typically 6) for each species within a plot had completed an event, representing the median value. Both approaches characterize central tendency for event timing across multiple individuals per plot. We used these data to determine changes in growing and reproductive season length, presenting data annually, and across years for the duration of species' active and reproductive periods. The start and end of the growing season were characterized by the mean across replicate plots for leaf emergence by the first species to leaf and for canopy senescence by the last species, respectively. Similarly, the start and end of the reproductive season were characterized by the date on which the first species flowered and seed maturation by the last species, respectively.

**Climate and microclimate.** Mean daily and mean annual temperature (MAT) and mean daily and total annual precipitation (TAP) were calculated on the basis of half hourly data from a meteorological station (HOBO, Onset, Inc.) at the field site (Extended Data Fig. 1). Growing degree day calculations were completed using data from infrared radiometers located within the experimental plots and a base temperature of  $0^{\circ}\text{C}$ . In each plot, the volumetric soil water content (SWC) was measured hourly at 10, 20, 40, 60 and 80 cm depth (EnviroSMART probe; Sentek Sensor Technologies). Daily means were calculated for SWC at the primary rooting depth (5–25 cm) by averaging the values for the sensors at 10 and 20 cm depth (Extended Data Fig. 2). We present data on mean spring air temperature across DOY 60–120, annual precipitation for the study site, and mean growing season (DOY 60–334) and autumn (DOY 244–304) SWC for the control plots (Fig. 1). Mean autumn SWC is also presented for experimental plots (Fig. 5).

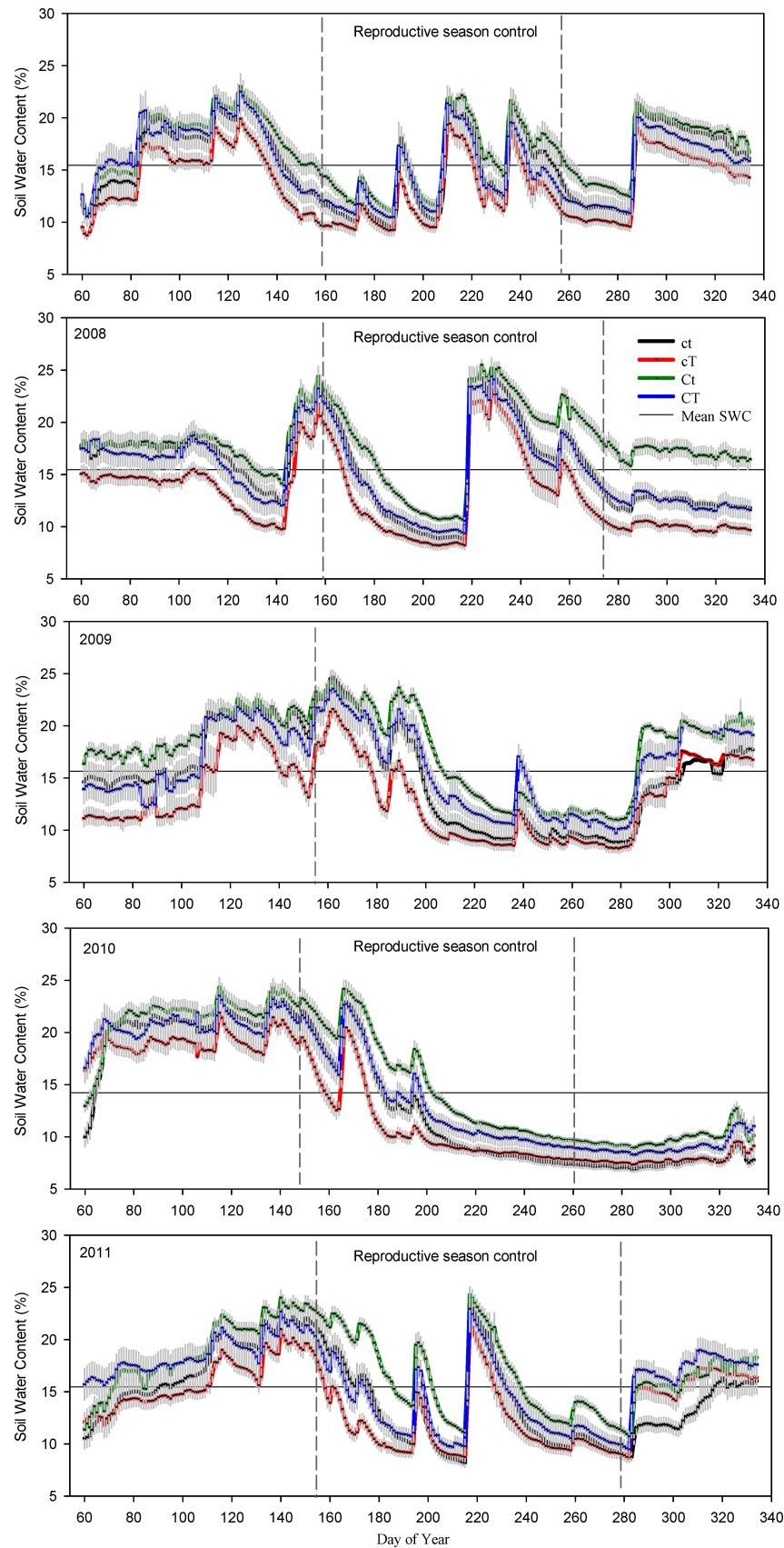
**Data analysis.** We analysed the data across years through ANOVA for each life cycle event, duration of active and reproductive periods, and autumn SWC, using Proc Mixed to test the main effects of temperature, elevated  $\text{CO}_2$ , species (if applicable) and year, and all their interactions (SAS version 9.2). For all ANOVAs, soil block based on the two soil types at the site was included as a random effect in the analyses and was not significant. Mixed model ANOVA allows for unequal variances and data were near a normal distribution. Analyses were done on untransformed data. Significant main effects and interactions ( $P < 0.05$ ) are reported on figures with complete ANOVA results reported in Extended Data Tables 1–3.



**Extended Data Figure 1 | Climate and warming effect for 2007–2011.**

**a, b,** Seasonal variation in precipitation and air temperature for the study site (**a**) and cumulative growing degree days (GDD) in control and warmed plots,

averaged across CO<sub>2</sub> levels (**b**) (means,  $n = 10$  plots). Mean annual temperature (MAT) and total annual precipitation (TAP) are listed each year.



**Extended Data Figure 2** | Seasonal variation in soil water content for 2007–2011. Values are means  $\pm$  1 s.e.m. for soil depth 5–25 cm ( $n = 5$  plots). Mean

annual soil water content (SWC) for control is represented by the horizontal grey line; vertical dashed lines show reproductive season timing for control.

Extended Data Table 1 | ANOVA results for the timing of annual life cycle events 2007–2011

Years 2007–2011	Leaf emergence		Flower production		Seed maturation		Canopy senescence	
Main effects	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>
Temp	<b>143.7*</b>	<b>&lt;0.001</b>	<b>42.0</b>	<b>&lt;0.001</b>	<b>12.5</b>	<b>&lt;0.001</b>	0.4	0.503
CO <sub>2</sub>	0.3	0.594	<b>7.5</b>	<b>0.007</b>	<b>17.7</b>	<b>&lt;0.001</b>	<b>9.5</b>	<b>0.002</b>
Species	<b>136.1</b>	<b>&lt;0.001</b>	<b>1639.5</b>	<b>&lt;0.001</b>	<b>1557.8</b>	<b>&lt;0.001</b>	<b>74.0</b>	<b>&lt;0.001</b>
Year	<b>50.3</b>	<b>&lt;0.001</b>	<b>42.2</b>	<b>&lt;0.001</b>	<b>9.7</b>	<b>&lt;0.001</b>	<b>118.8</b>	<b>&lt;0.001</b>
Temp x CO <sub>2</sub>	1.0	0.325	0.1	0.793	0.0	0.998	2.5	0.117
Temp x species	<b>5.1</b>	<b>&lt;0.001</b>	<b>5.1</b>	<b>&lt;0.001</b>	<b>3.4</b>	<b>0.006</b>	<b>3.6</b>	<b>0.003</b>
CO <sub>2</sub> x species	1.0	0.419	<b>3.4</b>	<b>0.005</b>	<b>4.0</b>	<b>0.002</b>	1.9	0.094
Temp x year	<b>28.0</b>	<b>&lt;0.001</b>	0.8	0.529	<b>6.0</b>	<b>&lt;0.001</b>	1.4	0.224
CO <sub>2</sub> x year	0.1	0.956	0.7	0.605	<b>6.1</b>	<b>&lt;0.001</b>	<b>5.8</b>	<b>&lt;0.001</b>
Species x year	<b>10.3</b>	<b>&lt;0.001</b>	<b>27.1</b>	<b>&lt;0.001</b>	<b>94.9</b>	<b>&lt;0.001</b>	<b>12.1</b>	<b>&lt;0.001</b>
Temp x CO <sub>2</sub> x species	0.4	0.859	2.2	0.060	0.3	0.928	1.5	0.191
Temp x CO <sub>2</sub> x year	0.2	0.905	1.1	0.342	0.8	0.556	0.6	0.685
Temp x species x year	<b>1.8</b>	<b>0.043</b>	1.5	0.106	<b>2.3</b>	<b>0.002</b>	1.3	0.154
CO <sub>2</sub> x species x year	0.7	0.824	1.6	0.054	<b>5.2</b>	<b>&lt;0.001</b>	1.5	0.072
Temp x CO <sub>2</sub> x species x year	1.1	0.377	1.7	0.053	1.1	0.381	<b>1.6</b>	<b>0.048</b>

\*Significant effects and interactions are in bold.



**Extended Data Table 2 | ANOVA results for the duration of species' active and reproductive periods**

Years 2007-2011	Active period duration		Reproductive period duration	
Main effects	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>
Temp	<b>20.1*</b>	<b>&lt;0.001</b>	<b>10.0</b>	<b>0.002</b>
CO <sub>2</sub>	<b>4.2</b>	<b>0.04</b>	0.3	0.59
Species	<b>103.9</b>	<b>&lt;0.001</b>	<b>84.6</b>	<b>&lt;0.001</b>
Year	<b>34.0</b>	<b>&lt;0.001</b>	<b>42.1</b>	<b>&lt;0.001</b>
Temp x CO <sub>2</sub>	0.1	0.72	0.0	0.98
Temp x species	<b>6.1</b>	<b>&lt;0.001</b>	1.3	0.28
CO <sub>2</sub> x species	0.9	0.48	0.3	0.91
Temp x year	<b>7.7</b>	<b>&lt;0.001</b>	<b>3.3</b>	<b>0.01</b>
CO <sub>2</sub> x year	2.3	0.08	<b>2.6</b>	<b>0.04</b>
Species x year	<b>10.3</b>	<b>&lt;0.001</b>	<b>36.2</b>	<b>&lt;0.001</b>
Temp x CO <sub>2</sub> x species	1.3	0.28	1.5	0.18
Temp x CO <sub>2</sub> x year	2.0	0.12	1.6	0.17
Temp x species x year	<b>2.7</b>	<b>0.001</b>	1.5	0.09
CO <sub>2</sub> x species x year	1.4	0.14	<b>2.4</b>	<b>0.002</b>
Temp x CO <sub>2</sub> x species x year	1.4	0.14	<b>2.4</b>	<b>0.005</b>

\* Significant effects and interactions are in bold.

Extended Data Table 3 | ANOVA results for autumn soil water content

Years 2007-2011	Autumn soil water content	
Main effects	<i>F</i>	<i>P</i>
Temp	<b>17.6*</b>	<b>&lt;0.001</b>
CO <sub>2</sub>	<b>57.5</b>	<b>&lt;0.001</b>
Year	<b>62.2</b>	<b>&lt;0.001</b>
Temp x CO <sub>2</sub>	2.9	0.09
Temp x year	<b>5.0</b>	<b>0.001</b>
CO <sub>2</sub> x year	1.5	0.21
Temp x CO <sub>2</sub> x year	0.2	0.94

\* Significant effects and interactions are in bold.

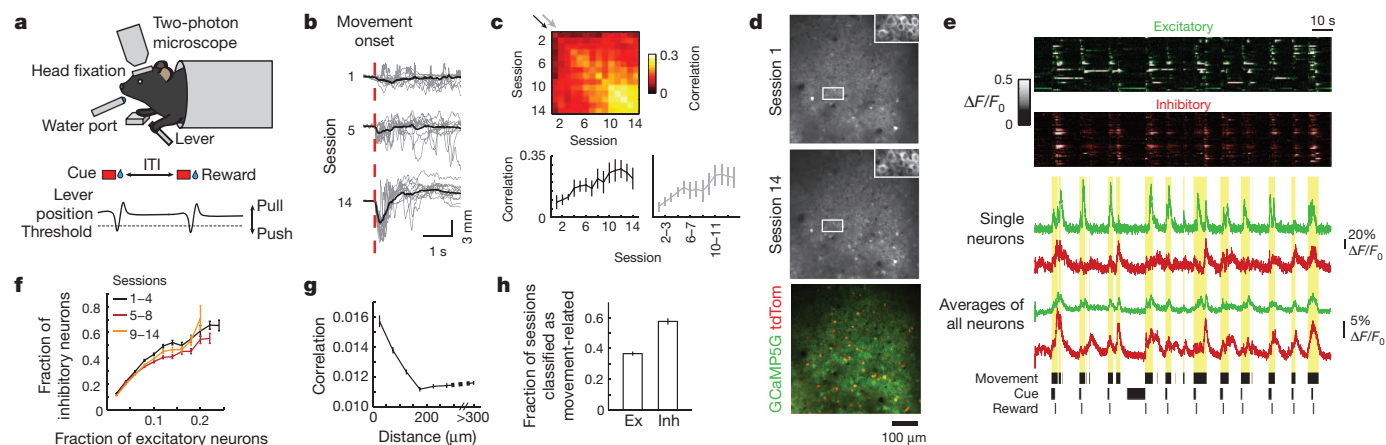
# Emergence of reproducible spatiotemporal activity during motor learning

Andrew J. Peters<sup>1</sup>, Simon X. Chen<sup>1</sup> & Takaki Komiyama<sup>1,2</sup>

The motor cortex is capable of reliably driving complex movements<sup>1,2</sup> yet exhibits considerable plasticity during motor learning<sup>3–10</sup>. These observations suggest that the fundamental relationship between motor cortex activity and movement may not be fixed but is instead shaped by learning; however, to what extent and how motor learning shapes this relationship are not fully understood. Here we addressed this issue by using *in vivo* two-photon calcium imaging<sup>11</sup> to monitor the activity of the same population of hundreds of layer 2/3 neurons while mice learned a forelimb lever-press task over two weeks. Excitatory and inhibitory neurons were identified by transgenic labelling<sup>12,13</sup>. Inhibitory neuron activity was relatively stable and balanced local excitatory neuron activity on a movement-by-movement basis, whereas excitatory neuron activity showed higher dynamism during the initial phase of learning. The dynamics of excitatory neurons during the initial phase involved the expansion of the movement-related population which explored various activity patterns even during similar movements. This was followed by a refinement into a smaller population exhibiting reproducible spatiotemporal sequences of activity. This pattern of activity associated with the learned movement was

unique to expert animals and not observed during similar movements made during the naive phase, and the relationship between neuronal activity and individual movements became more consistent with learning. These changes in population activity coincided with a transient increase in dendritic spine turnover in these neurons. Our results indicate that a novel and reproducible activity–movement relationship develops as a result of motor learning, and we speculate that synaptic plasticity within the motor cortex underlies the emergence of reproducible spatiotemporal activity patterns for learned movements. These results underscore the profound influence of learning on the way that the cortex produces movements.

We developed a cued lever-press task performed by mice under a two-photon microscope (Fig. 1a), similar to other recently reported tasks<sup>14,15</sup>. Briefly, a lever press beyond the set threshold during an auditory cue was rewarded with water (Methods). Mice were trained with this task daily for 2 weeks ( $n = 10$ ). Even though mice achieved a reward in most trials, the timing of their behaviour improved in later sessions (Extended Data Fig. 1). Furthermore, lever movements on individual trials became more stereotyped over time (Fig. 1b). The reproducibility of movement



**Figure 1 | Lever-press task and chronic calcium imaging of excitatory and inhibitory populations.** **a**, Task schematic. **b**, Lever movement traces in rewarded trials from one mouse. Grey, 10 individual trials; black, average of all trials; red dotted line, movement onset. **c**, Top: median pairwise correlation coefficients of rewarded movements on individual trials over 3 s, averaged across animals. Bottom: pairwise movement correlation on individual trials within and across sessions corresponding to the black and grey arrows indicated on the top, respectively. Individual movements became more similar within ( $r = 0.35$ ,  $P < 0.001$ ) and across ( $r = 0.37$ ,  $P < 0.001$ ) sessions. See Methods for sample size. **d**, Top and middle: GCaMP5G expression in layer 2/3 neurons imaged 2 weeks apart. Insets: magnified images of outlined areas. Bottom: merge of tdTomato expressed in all inhibitory neurons (red) and GCaMP5G (green). **e**, Top: activity of all simultaneously imaged movement-related 38 excitatory (green) and 42 inhibitory (red) neurons from one animal. Each row represents a neuron. Middle:  $\Delta F/F_0$  traces from one neuron each and

average  $\Delta F/F_0$  of all imaged neurons of each type (152 excitatory and 77 inhibitory). Bottom: task-related events; yellow shading indicates movement epochs. **f**, Fractions of active inhibitory neurons and excitatory neurons during rewarded movements are correlated on a movement-by-movement basis ( $r = 0.63$ – $0.67$ ,  $P < 0.001$ ). This relationship is stable throughout learning ( $P = 0.92$ , one-way ANOVA comparison of median excitatory/inhibitory ratios). **g**, Pairwise correlation coefficients between inhibitory and excitatory neuron activity decrease with distance ( $P < 0.001$ , comparison between pairs within 150  $\mu\text{m}$  and all other pairs, Wilcoxon rank sum test,  $n = 653,046$  excitatory–inhibitory pairs total). **h**, Individual movement-related inhibitory neurons are classified on more sessions, showing that excitatory neurons are on average more dynamic than inhibitory neurons ( $P < 0.001$ , Wilcoxon rank sum test,  $n = 473$  and 231 movement-related excitatory and inhibitory neurons, respectively). All error bars are s.e.m.

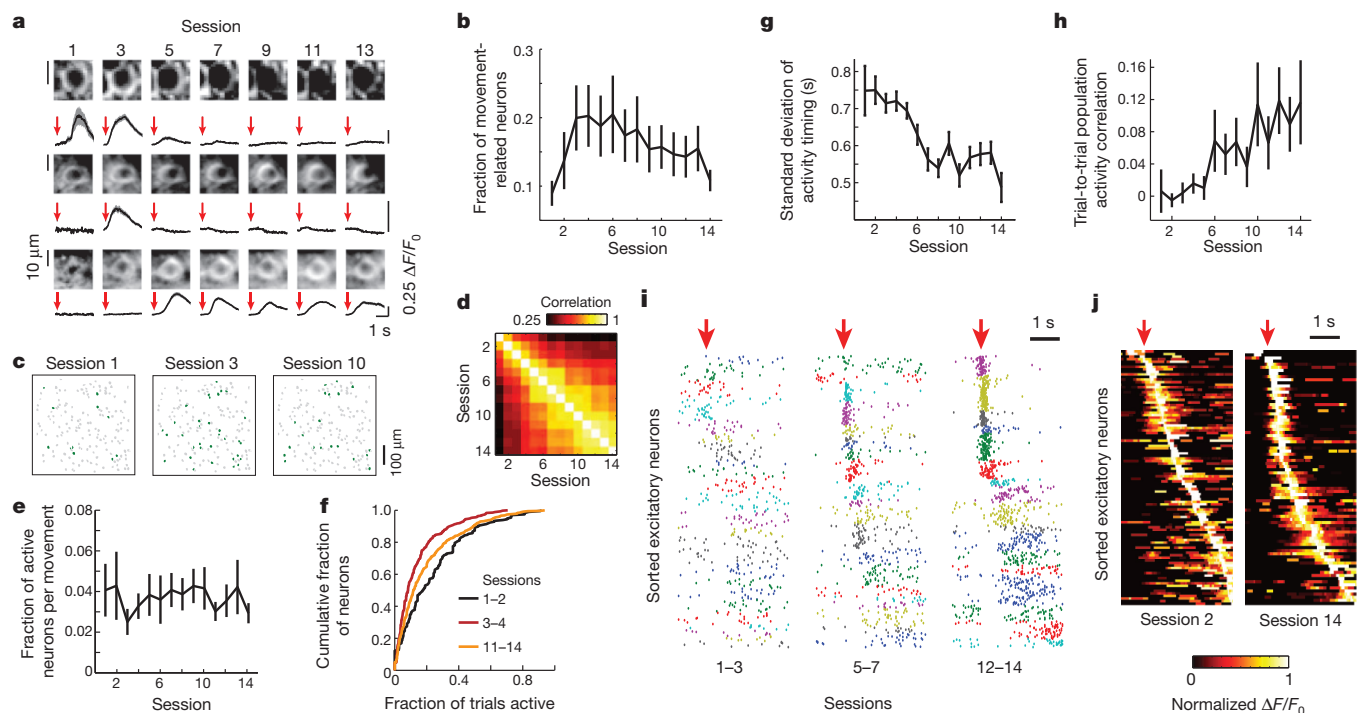
<sup>1</sup>Neurobiology Section, Center for Neural Circuits and Behavior, and Department of Neurosciences, University of California, San Diego, La Jolla, California 92093, USA. <sup>2</sup>JST, PRESTO, University of California, San Diego, La Jolla, California 92093, USA.

kinematics was evident in higher correlation of rewarded movements on individual trials within and across later sessions (Fig. 1c). The motor cortex is necessary for this task, as lesions before training prevented the emergence of movement stereotypy, and acute inactivation by pharmacology or optogenetics impaired task performance (Extended Data Fig. 2).

To identify how the activity of motor cortex neuronal ensembles is modified during this learning, we combined the lever-press task with chronic two-photon calcium imaging<sup>8,16</sup>. In this study we focused on neurons in layer 2/3, the major input layer capable of driving deeper layer neurons to produce motor cortex outputs<sup>17,18</sup>. Before training, we injected an adeno-associated virus encoding the  $\text{Ca}^{2+}$  indicator GCaMP5G<sup>19</sup> into the right forelimb area of the motor cortex to express GCaMP5G in all neuron types. Optogenetic stimulation of this area evoked forelimb movements (Extended Data Fig. 3). GCaMP5G fluorescence reported spiking activity with high temporal precision (jitter of spikes and calcium events =  $7.1 \pm 41.4$  ms, median  $\pm$  standard deviation (s.d.); Extended Data Fig. 4). We used transgenic mice that express tdTomato in all GABAergic inhibitory neurons (*Gad2-IRES-Cre* (ref. 12); *Rosa-LSL-tdTomato* (ref. 13)) to identify excitatory (only expressing GCaMP5G) and inhibitory (expressing both tdTomato and GCaMP5G) neurons. Two weeks after surgery, hundreds of neurons were imaged through a chronic window. A total of  $202 \pm 18$  (mean  $\pm$  standard error of the mean (s.e.m.)) neurons were imaged in each animal, with  $20.9 \pm 2.6\%$  being inhibitory, consistent with the composition of the cortex<sup>20</sup>.

We imaged the activity of the same population of excitatory and inhibitory neurons over the course of 2 weeks while mice simultaneously learned and performed the lever-press task (Fig. 1d,  $n = 7$  mice). High correlations of neural activity and lever-press movements were evident both at the level of single neurons and population average in both excitatory and inhibitory neurons (Fig. 1e). A large fraction of imaged neurons showed significantly more activity during lever-press movements and were thus considered movement-related ( $51.4 \pm 5.7\%$  of imaged neurons,  $44.0 \pm 6.2\%$  of excitatory and  $78.7 \pm 5.7\%$  of inhibitory, were classified as movement-related in at least one session; Methods). Movement-related neurons did not show obvious spatial clustering (Extended Data Fig. 5).

We investigated the relationship between excitatory and inhibitory populations. We found a positive correlation between the fraction of active inhibitory and excitatory neurons on a movement-by-movement basis; during movements that activated a larger fraction of excitatory neurons, a larger fraction of inhibitory neurons were also active. This relationship of excitatory and inhibitory activity remained constant for the entire 2 weeks of imaging (Fig. 1f). Individual inhibitory neurons were particularly correlated with nearby excitatory neurons within  $150 \mu\text{m}$ , consistent with their connectivity<sup>21,22</sup> (Fig. 1g). This local matching of excitatory and inhibitory activity probably provides a basis for the balance between excitatory and inhibitory inputs to individual neurons observed in the cortex (reviewed in ref. 20). Even though the ratio of excitatory and inhibitory



**Figure 2 | Dynamics of spatiotemporal activity of excitatory neurons during learning.** **a**, Dynamics of three excitatory neurons. Top rows: images of neurons confirming reliable identification. Bottom rows: black, mean  $\Delta F/F_0$ ; grey, s.e.m.; arrow, movement onset. **b**, Mean fraction of excitatory neurons classified as movement-related in each session (increase over sessions 1–3,  $P < 0.01$ ; decrease over sessions 4–14,  $P < 0.01$ , Methods). **c**, Dynamic population of excitatory neurons from one mouse. Green, movement-related excitatory neurons; grey, non-classified excitatory neurons. **d**, Correlation of population activity of all excitatory neurons during rewarded movements across sessions (Methods). The population of movement-related excitatory neurons became more stable in later sessions ( $P < 0.001$ , session 1–4 pairs versus session 10–14 pairs, Wilcoxon rank sum test). **e**, Average fraction of excitatory neurons that are active in each individual movement out of all excitatory neurons remains stable throughout learning ( $r = 0.00$ ,  $P = 0.98$ ). **f**, Cumulative distribution of fraction of trials in which each movement-related excitatory neuron is active. In sessions 3–4, individual movement-related neurons are active in fewer trials compared to sessions 1–2 or 11–14 ( $P < 0.001$ ,

Kolmogorov–Smirnov test). **g**, Standard deviation of the timing of activity onsets for movement-related excitatory neurons decreases over sessions ( $r = -0.24$ ,  $P < 0.001$ ). Neurons that were active in less than five trials of a given session were excluded from this analysis. **h**, Pairwise trial-to-trial correlation of temporal population activity vectors increases with learning ( $r = 0.43$ ,  $P < 0.001$ ). Temporal population activity vector was a concatenation of the activity traces of all movement-related neurons and thus maintained temporal information within each movement. **i**, Activity onsets of excitatory neurons from one animal that are movement-related and active in at least 10% of trials on the sessions indicated. Arrow, movement onsets; colours, individual neurons sorted according to their preferred timing. Note that same colours across sessions are not necessarily the same neurons. **j**, Maximum-normalized average activity from all movement-related neurons from all animals in session 2 (left, 106 neurons) and session 14 (right, 84 neurons) aligned to movement onset (arrow). Activity timing is refined over time, shown by narrower peaks and lower background in session 14, and shifts towards movement onset. See Methods for sample size. All error bars are s.e.m.

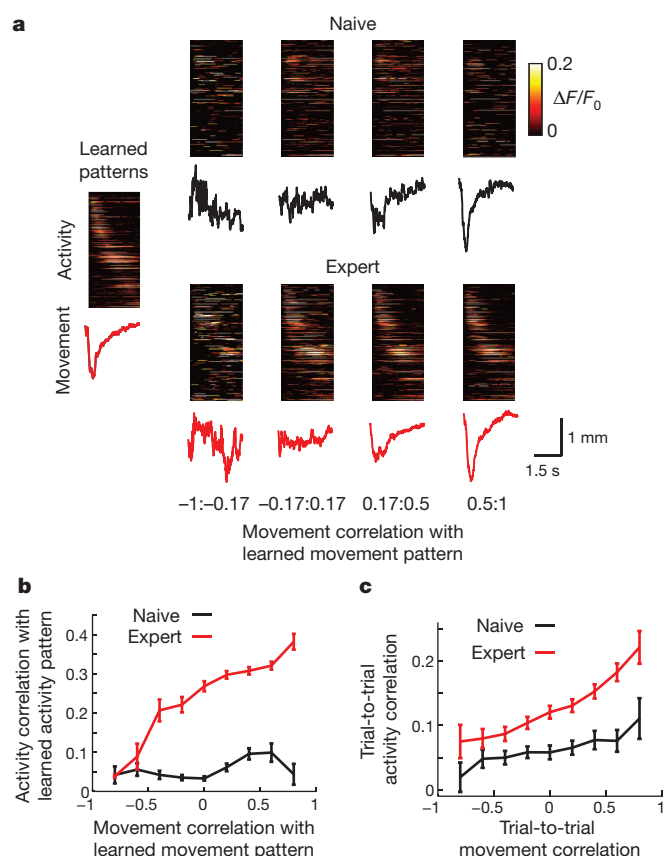


activity on a trial-to-trial basis was stable, the identity of movement-related neurons was dynamic. In particular, excitatory neurons on average had a higher degree of turnover than the inhibitory population, indicating that the excitatory population is more dynamic during learning (Fig. 1h and Extended Data Fig. 6a). We therefore focused on excitatory neurons for the following analyses.

Many excitatory neurons were transiently movement-related (Fig. 2a). In the initial phase of learning, a large fraction of excitatory neurons developed movement-related activity, resulting in a marked increase of the movement-related population (Fig. 2b, c). After this initial expansion, the fraction of movement-related neurons decreased gradually through the remaining course of the experiment (Fig. 2b, c and Extended Data Fig. 6b). This resulted in a smaller and more stable population of movement-related neurons towards the end of learning (Fig. 2d). The expansion and refinement was not seen during spontaneous movements without training (Extended Data Fig. 6c). Despite these changes in the ensemble of movement-related neurons during learning, the average fraction of excitatory neurons active on each trial remained stable (Fig. 2e). This constant level of activity was maintained despite a changing size of movement-related populations because of the corresponding shift in the frequency of activity in individual neurons (Fig. 2f). Various combinations of excitatory neurons were therefore used within the motor cortex during the initial phase of learning, followed by a refinement of the population to form a stable network associated with the learned movement.

We next examined the timing of activity of individual neurons. As a population, the activity of movement-related excitatory neurons diverged from baseline at 105 ms before the movement onset and continued throughout the duration of movements (Extended Data Fig. 6d). During the first few sessions of learning, movement-related excitatory neurons showed variable timing of activity on individual trials relative to the movement onset. Conversely, movement-related neurons in later sessions showed reproducible activity timing relative to movement onset (Fig. 2g and Extended Data Fig. 6e). As a result, the temporal activity pattern became progressively more stable during learning (Fig. 2h). This activity sequence tiled the entire duration of movement (Fig. 2i, j and Extended Data Fig. 6f). Furthermore, the population activity shifted towards the beginning of movements over the course of the experiment (Fig. 2j and Extended Data Fig. 6g).

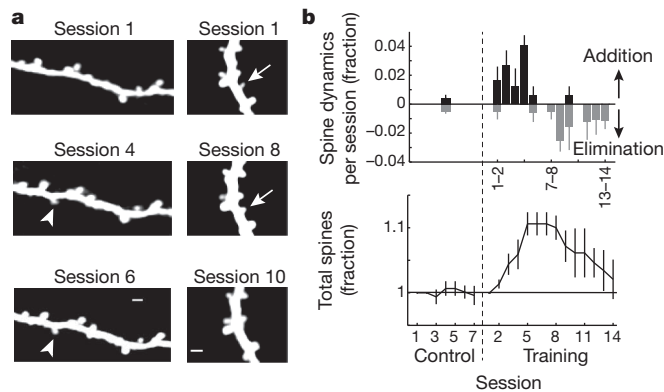
The observed stabilization of motor cortex activity may result from the selection of a particular activity–movement pair out of many that are explored during initial learning. In this case, activity during the learned movement and activity during similar movements made early in learning should resemble each other. Alternatively, the ‘learned’ activity pattern may be unique to the expert stage after learning. To distinguish between these possibilities, we evaluated the relationship between movement and neural activity. We defined the ‘learned’ patterns of activity and movement as the averages of a randomly chosen 50% of trials from the expert sessions (sessions 10–14, Methods). When the remaining 50% of trials from the expert sessions were sorted according to the similarity of the movement in each trial with the learned movement pattern, a clear relationship between movement and activity became evident—the similarity of activity to the learned activity pattern increased with the similarity of movements to the learned movement (Fig. 3a, b). Remarkably, this relationship was much weaker when trials from naive sessions (sessions 1–3) were considered. Regardless of the similarity of movement to the learned movement pattern, the similarity of activity to the learned activity pattern was consistently low in naive sessions (Fig. 3a, b). In other words, the learned activity pattern was reproducibly observed only when the expert animals made the learned movement, whereas similar movements made in naive sessions were accompanied by very different activity patterns. Furthermore, the general relationship between activity and movement in pairs of trials became more consistent after learning, whereas activity in naive animals was more variable regardless of movement similarity (Fig. 3c). These analyses of learning-related changes in population activity were performed using the entire movement periods. However,



**Figure 3 | Learning-related emergence of reproducible spatiotemporal activity.** **a**, Heat maps show mean activity of movement-related excitatory neurons classified during expert sessions of all animals aligned to movement onset (left edge). Traces show mean lever movements. Trials are binned according to the correlations of the movements on those trials with the learned movement pattern (Methods). Left, learned patterns; top, naive sessions (sessions 1–3); bottom, expert sessions (sessions 10–14). **b**, Correlation of trial activity with the learned activity pattern increases with the correlation of trial movement with the learned movement pattern in expert sessions. Movements similar to the learned movement pattern but made in naive sessions display activity very different from the learned activity pattern ( $P = 0.83, 0.35$  and  $<0.001$  in the bins 1, 2 and 3–9, respectively, Wilcoxon rank sum test). **c**, Pairwise trial-to-trial correlation of temporal population activity vectors (defined as in Fig. 2h) plotted as a function of movement correlation on those trials. A stronger relationship between population activity and movement emerges during learning ( $P = 0.38, 0.18, 0.04, 0.002, <0.001$  and  $0.02$  in bins 1, 2, 3, 4, 5–8 and 9, respectively, Wilcoxon rank sum test). See Methods for sample size. All error bars are s.e.m.

the results were similar when only the periods from movement onset to reward were considered (Extended Data Fig. 7).

The plasticity of population activity described above could simply reflect changes in other brain areas providing inputs to these neurons. However, synaptic plasticity within the motor cortex could also contribute to the changes in population activity. To test whether learning of this lever-press task induces synaptic plasticity in layer 2/3 of the motor cortex, we labelled sparse subsets of layer 2/3 neurons and chronically imaged spines on the same dendritic branches of excitatory neurons over the course of learning ( $n = 191$  spines in 3 mice). Imaging was performed immediately before each training session in awake animals. We observed the formation of a number of dendritic spines during the initial sessions of training, followed by the elimination of some of the spines that were present at the beginning of the experiment. Most (95%, 19 out of 20) of the spines that formed during training persisted for the entire 2 weeks. At the population level, these changes resulted in a transient 10% increase in the density of spines followed by return to the baseline



**Figure 4 | Learning-related plasticity of dendritic spines.** **a**, Example layer 2/3 excitatory neuron dendrites imaged in awake mice throughout learning. Arrowheads, added spine; arrows, eliminated spine; scale bars, 2  $\mu$ m. **b**, Summary of spine dynamics in trained and control animals. Top: spine additions (black) and eliminations (grey) in each session. For control animals, data from all sessions are combined. Bottom: total spine number across sessions normalized to session 1 in each condition. Spine density transiently increases during learning ( $P < 0.001$ , control versus training sessions 4–7, Wilcoxon rank sum test).  $n = 191$  spines in 3 mice. All error bars are s.e.m.

(Fig. 4). These results are analogous to previous reports in motor cortex layer 5 neurons during different motor learning tasks<sup>3,4</sup>. Spines were largely stable in a separate group of animals that did not undergo training but otherwise were treated identically including water restriction and head fixation (Fig. 4b). The spine density was also stable in the hind-limb area in the motor cortex during learning (Extended Data Fig. 8). These results indicate that our lever-press task induces area-specific reorganizations of excitatory synapses onto layer 2/3 neurons during learning.

Previous studies suggested a relatively stable tuning and population code of motor cortex neurons in well-trained animals<sup>23,24</sup>. However, our understanding is quite limited as to the changes in the ensemble activity pattern during the transition from naive to expert stages. Our results indicate that the relationship between movements and activity is initially inconsistent (that is, degenerate), and the early days of learning involve the expansion and exploration of various movement-related activity in the motor cortex. An increased variability of single-neuron activity in the motor cortex has been observed during learning of visuomotor adaptation<sup>25</sup> and a brain–machine interface task<sup>26</sup>. Such trial-to-trial variability has been proposed to provide the basis for exploration of possible network states and facilitate learning<sup>10</sup>. Our results directly demonstrate such an exploration during initial learning at the population level. After this period of high variability, the activity–movement degeneracy is reduced and a reproducible temporal sequence of activity emerges in a stable population of excitatory neurons (Extended Data Fig. 9). Such spatiotemporal activity may orchestrate the temporal dynamics of the learned movement. Reproducible temporal patterns of population activity during learned movements are proposed to be generated by internal connections within the motor cortex<sup>27,28</sup>. We note that our results do not provide a causal link between local synaptic plasticity and changes in population activity. Nevertheless, we show that these processes occur during motor learning on similar timescales, which supports the notion that local synaptic plasticity may generate a circuit to reproduce a particular spatiotemporal activity pattern. These new circuits may be more efficient in driving movements, which could underlie the lower metabolic activity in the motor cortex observed during execution of well-practiced movements<sup>29,30</sup>. Our study provides a glimpse of the emergence of population activity patterns for learned movements.

## METHODS SUMMARY

Surgeries were performed to inject viruses in the right forelimb area of the motor cortex and implant a chronic window and head plate. For functional imaging, AAV2/1-Syn-GCaMP5G was injected in *Gad2-IRES-Cre*; *Rosa-LSL-tdTomato* mice; for

structural imaging, a mixture of AAV2/1-CAG-FLEX-EGFP and AAV2/1-CMV-PI-Cre was injected in C57BL/6 mice. Imaging and behaviour experiments started at least 2 weeks after surgery and mice were at least 8 weeks of age. Imaging was performed in awake mice during (functional imaging) or right before (structural imaging) each behavioural session.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 August 2013; accepted 7 March 2014.

Published online 4 May 2014.

- Graziano, M. S. A. *The Intelligent Movement Machine* 1st edn (Oxford, 2009).
- Harrison, T. C., Ayling, O. G. & Murphy, T. H. Distinct cortical circuit mechanisms for complex forelimb movement and motor map topography. *Neuron* **74**, 397–409 (2012).
- Xu, T. et al. Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* **462**, 915–919 (2009).
- Yang, G., Pan, F. & Gan, W. B. Stably maintained dendritic spines are associated with lifelong memories. *Nature* **462**, 920–924 (2009).
- Nudo, R. J., Milliken, G. W., Jenkins, W. M. & Merzenich, M. M. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* **16**, 785–807 (1996).
- Rioult-Pedotti, M. S., Friedman, D. & Donoghue, J. P. Learning-induced LTP in neocortex. *Science* **290**, 533–536 (2000).
- Komiyama, T. et al. Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464**, 1182–1186 (2010).
- Huber, D. et al. Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* **484**, 473–478 (2012).
- Sanes, J. N. & Donoghue, J. P. Plasticity and primary motor cortex. *Annu. Rev. Neurosci.* **23**, 393–415 (2000).
- Rokni, U., Richardson, A. G., Bizzi, E. & Seung, H. S. Motor learning with unstable neural representations. *Neuron* **54**, 653–666 (2007).
- Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl Acad. Sci. USA* **100**, 7319–7324 (2003).
- Taniguchi, H. et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995–1013 (2011).
- Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
- Isomura, Y., Harukuni, R., Takekawa, T., Aizawa, H. & Fukai, T. Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. *Nature Neurosci.* **12**, 1586–1593 (2009).
- Hira, R. et al. Spatiotemporal dynamics of functional clusters of neurons in the mouse motor cortex during a voluntary movement. *J. Neurosci.* **33**, 1377–1390 (2013).
- Kato, H. K., Chu, M. W., Isaacson, J. S. & Komiyama, T. Dynamic sensory representations in the olfactory bulb: modulation by wakefulness and experience. *Neuron* **76**, 962–975 (2012).
- Weiler, N., Wood, L., Yu, J., Solla, S. A. & Shepherd, G. M. Top-down laminar organization of the excitatory network in motor cortex. *Nature Neurosci.* **11**, 360–366 (2008).
- Kaneko, T., Cho, R., Li, Y., Nomura, S. & Mizuno, N. Predominant information transfer from layer III pyramidal neurons to corticospinal neurons. *J. Comp. Neurol.* **423**, 52–65 (2000).
- Akerboom, J. et al. Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
- Isaacson, J. S. & Scanziani, M. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).
- Kwan, A. C. & Dan, Y. Dissection of cortical microcircuits by single-neuron stimulation *in vivo*. *Curr. Biol.* **22**, 1459–1467 (2012).
- Kätzel, D., Zemelman, B. V., Buettner, C., Wolfel, M. & Miesenböck, G. The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nature Neurosci.* **14**, 100–107 (2011).
- Chestek, C. A. et al. Single-neuron stability during repeated reaching in macaque premotor cortex. *J. Neurosci.* **27**, 10742–10750 (2007).
- Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419 (1986).
- Mandelblat-Cerf, Y., Paz, R. & Vaadia, E. Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. *J. Neurosci.* **29**, 15053–15062 (2009).
- Arduni, P. J., Fregnac, Y., Shulz, D. E. & Ego-Stengel, V. “Master” neurons induced by operant conditioning in rat motor cortex during a brain-machine interface task. *J. Neurosci.* **33**, 8308–8320 (2013).
- Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).
- Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**, 394–399 (2010).
- Krings, T. et al. Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci. Lett.* **278**, 189–193 (2000).

30. Picard, N., Matsuzaka, Y. & Strick, P. L. Extended practice of a motor skill is associated with reduced metabolic activity in M1. *Nature Neurosci.* **16**, 1340–1347 (2013).

**Acknowledgements** We thank A. Kim and S. Kalina for technical assistance; L. L. Looger, J. Akerboom, D. S. Kim and the GENIE Project at Janelia Farm for making GCaMP available; E. Kyubwa and J. Keller for help with task development; A. D. Lien and M. Caudill for help with two-photon guided recordings; and M. Long, R. Malinow, G. Murphy, M. Scanziani and members of the Komiyama laboratory for comments and discussions. This work was supported by grants from Japan Science and Technology Agency (PRESTO), Pew Charitable Trusts, Alfred P. Sloan Foundation, David & Lucile Packard Foundation, Human Frontier Science Program and New York Stem Cell Foundation to T.K. A.J.P. is supported by the Neuroplasticity of Aging Training Grant

(AG000216). S.X.C. is a Human Frontier Science Program postdoctoral fellow. T.K. is a NYSCF-Robertson Investigator.

**Author Contributions** A.J.P. and T.K. conceived the project. Dendritic spine imaging and optogenetic silencing experiments were performed by S.X.C. and analysed by S.X.C. and T.K. All other experiments were performed by A.J.P. and analysed by A.J.P. and T.K. A.J.P. and T.K. wrote the manuscript with input from S.X.C.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.K. ([tkomiyama@ucsd.edu](mailto:tkomiyama@ucsd.edu)).

## METHODS

**Animals.** All procedures were in accordance with protocols approved by the UCSD Institutional Animal Care and Use Committee and guidelines of the National Institute of Health. Mice (calcium imaging: cross between *Gad2-IRES-Cre* (ref. 12) and *Rosa26-CAG-LSL-tdTomato* (ref. 13), Jackson Laboratories; structural imaging: C57BL/6, Charles River Laboratory) were group housed in disposable plastic cages with standard bedding in a room with a reversed light cycle (12 h–12 h). After surgery, animals were singly housed. Experiments were performed during the dark period. No randomization was used, but internal controls were included whenever possible.

**Surgery.** Adult mice (6 weeks or older, male and female) were anaesthetized with isoflurane and injected with dexamethasone ( $2 \text{ mg kg}^{-1}$ ) and baytril ( $10 \text{ mg kg}^{-1}$ ) intramuscularly to prevent brain swelling and infection. A custom head-plate was glued to the skull and craniotomy ( $\sim 2 \text{ mm}$  diameter) was performed as described<sup>31</sup> over the right caudal forelimb area. Viruses (UPenn Vector Core; calcium imaging: AAV2/1-syn-GCaMP5G diluted 1:1–3 in saline, 5–6 sites; structural imaging: AAV2/1-CAG-FLEX-EGFP (1:1) and AAV2/1-CMV-PI-Cre (1:5,000) diluted in saline, 3 sites) were injected in the caudal forelimb area of the motor cortex around the coordinate of  $300 \mu\text{m}$  anterior and  $1,500 \mu\text{m}$  lateral from bregma, according to previous microstimulation experiments<sup>15,32–36</sup>. We confirmed that optogenetic stimulation of this area evoked forelimb movements (Extended Data Fig. 3). This area also falls near the border of the abduction and adduction areas defined by ref. 2. For control experiments targeting the hindlimb area of the motor cortex, viruses were injected around the coordinate of  $1,500 \mu\text{m}$  posterior and  $1,500 \mu\text{m}$  lateral from bregma, according to previous microstimulation experiments<sup>32,33,35,36</sup>. Each injection consisted of  $\sim 20 \text{ nl}$  at a depth of  $\sim 250 \mu\text{m}$  from the pia injected over 2–4 min and each injection site was separated by  $\sim 500 \mu\text{m}$ . Pipettes were left in the brain for 3 min after injection to avoid backflow. After virus injections, a chronic imaging window was implanted consisting of a glass plug glued onto a larger glass base. The edges between the glass plug and the skull were filled with 1.5% agarose and the window was secured using dental acrylic. Buprenorphine ( $0.1 \text{ mg kg}^{-1}$ ) was injected subcutaneously at the end of surgery.

**Behaviour.** Three days after surgery, mice were water-restricted at 1 ml per day. After  $\sim 14$  days of water restriction, mice were trained daily for 14 days while two-photon imaging was applied simultaneously. The lever was built using a piezoelectric flexible force transducer (LCL-113G, Omega Engineering) attached to a 1/16-mm-thick brass rod. The voltage from the force transducer, which is proportional to the lever position, was continuously recorded using a data acquisition device (Labjack) and software (Ephus, MATLAB, Mathworks) working with custom software running on LabVIEW (National Instruments) which monitored threshold crossing. The behavioural setup was controlled by software (Dispatch, Z. Mainen and C. Brody) running on MATLAB communicating with a real-time system (RTLinux). A 6-kHz tone marked a period during which lever-press was rewarded with water ( $\sim 8 \mu\text{l}$  per trial) paired with a 500-ms, 12-kHz tone, followed by an intertrial interval (variable duration, see below). Lever-press was defined as crossing of two thresholds ( $\sim 1.5 \text{ mm}$  and  $\sim 3 \text{ mm}$  below the resting position) within 200 ms. The 3-mm threshold defined the displacement required, and the 1.5-mm threshold ensured that the mouse did not hold the lever near the lower threshold. Failure to press during the cue period triggered a loud white noise and an intertrial interval. Lever presses during intertrial intervals were neither rewarded nor punished. The cue period was decreased during the first two sessions from 30 s to 10 s. The reward period was reduced during the first three sessions from 2 s to 0.4 s. The intertrial interval was increased over the first three sessions from 2–4 s to 8–12 s. Each session lasted 20–30 min and 100–200 trials until terminated when mice stopped performing or consumed 1 ml of water. Experiments lasted for 11–14 sessions. One mouse failed to learn the task and was thus excluded from all analyses.

**Lesion.** Motor cortex lesions were performed under isoflurane anaesthesia. After craniotomy was performed as above, cortical tissue was aspirated using a glass Pasteur pipette connected to vacuum. Care was taken to avoid damaging the underlying white matter. After the lesion, the cavity was filled with Surgifoam (Johnson & Johnson), KWIK-CAST (World Precision Instruments) and then with a layer of dental acrylic. Mice were allowed to recover for 3 days after the surgery and then placed on water restriction. Behavioural training started 2 weeks after lesion. The extent of lesion was determined for each mouse with post hoc histology.

**Muscimol inactivation.** Mice with imaging windows were first trained with the task for 7–14 days. On the day of inactivation, the imaging window was removed and  $\sim 70 \text{ nl}$  of muscimol ( $5 \mu\text{g}/\mu\text{l}$  in cortex buffer) was injected over 2–3 min in the centre of the craniotomy at the depth of  $300 \mu\text{m}$ , under isoflurane anaesthesia. The craniotomy was then sealed and mice were allowed to recover in their home cage on a heating pad for 60 min before behavioural experiments. For control injections, muscimol was injected into the barrel cortex, using the coordinate of 1.4 mm posterior and 3.1 mm lateral from bregma.

**Optogenetic inactivation.** Surgery was performed on PV-Cre mice as above to inject AAV2/1-CAG-Flex-ChR2-tdTomato around the forelimb area of the motor cortex at five sites.  $80 \text{ nl}$  was injected at each site at each depth of  $400 \mu\text{m}$  and  $800 \mu\text{m}$ . Behavioural training started 3 weeks after surgery. After seven sessions of daily training, the cortical area was inactivated in 20% of trials by activating PV neurons<sup>37</sup> by blue light from an LED ( $\sim 40 \text{ mW}$ , 470 nm, Doric Lenses) delivered directly onto the centre of the craniotomy covered with a chronic glass window. Blue light was delivered starting from 1–2 s before the cue period until the end of the cue period (that is, reward delivery or time out). Blue light delivery was performed in seven successive sessions. In the last two sessions, the glass window was covered with curable silicone (KWIK-CAST, World Precision Instruments) and these served as control sessions.

**Optogenetic microstimulation.** Surgery was performed on C57BL/6 wild-type mice as above to inject AAV2/1-CAG-ChR2-Venus around the forelimb area of the motor cortex at 9 sites.  $20 \text{ nl}$  was injected at each site at the depth of  $300 \mu\text{m}$ . Three weeks after surgery, mice were head-fixed. Blue light from an LED ( $\sim 40 \text{ mW}$ , 470 nm, Doric Lenses) was delivered directly onto the centre of the craniotomy covered with a chronic glass window for 1 s per trial with the intertrial interval of 8–12 s. Stimulation was performed without anaesthesia. The effect of the light stimulus on forelimb movements was quantified by manually identifying movie frames in which the forelimb movements were observed on videos cropped so that the optical stimulation was not visible. Pre-light periods were defined as movie frames lasting from 2 to 1 s before light onset. Scoring was done by two individuals independently and trials which differed between the scorers were excluded (1/79 light trials, 3/79 pre-light trials).

**Imaging.** Imaging was conducted with a commercial two-photon microscope (Bioscope, Thorlabs) running Scanimage using a  $16\times$  objective (NIKON) with excitation at  $925 \text{ nm}$  (Ti-Sa laser, Newport). Imaging was always conducted in awake animals. For calcium imaging, images ( $512 \times 512$  pixels covering  $472 \times 508 \mu\text{m}$ ) were recorded at approximately 28 Hz in continuous segments about 2 min long each with inter-segment intervals of 12 s. The trials that overlapped with the intervals were discarded. Signals for the start of each trial were also recorded, which were used to align images and behaviour data. Slow drifts in imaging field were manually corrected using reference images. For structural imaging, stacks of image planes ( $512 \times 512$  pixels covering  $94 \times 104 \mu\text{m}$ ) were acquired at approximately 28 Hz, 20 frames per plane, 80–120 planes per animal with a z-axis step size of  $1.0 \mu\text{m}$  between planes.

**Movement analysis.** Movement bouts were identified in the lever displacement traces (voltage recordings from the force transducer) that were down-sampled from 10 kHz to 1 kHz and then filtered (4-pole 10 Hz low-pass Butterworth). The velocity of the lever was then determined by smoothing the difference of consecutive points with a moving average window of 5 ms. The envelope of the velocity was then extracted using a Hilbert transform, and movement bouts were defined by the envelope crossing a threshold ( $4.9 \text{ mm per second}$ ). Each movement bout was extended by 75 ms, bouts separated by less than 500 ms were considered continuous, and then the start and end times were fine-tuned as follows. The start time was defined by finding when the lever position crossed a threshold exceeding the resting period before the movement, and the end time was defined by finding when the lever position went below a threshold defined by the resting period following the movement. Thresholds were the resting position plus the 99th percentile of noise distribution defined as the difference between the Butterworth smoothed trace and the original trace. These processes were chosen empirically based on visual inspection. For trial-based analyses, the trials in which animals were moving the lever at the onset of cue were excluded.

**Image analysis.** For calcium imaging data, lateral motion was corrected using full-frame cross-correlation image alignment (Turboreg<sup>38</sup> plugin for ImageJ). Motion within each frame was negligible due to the fast frame rate. Regions of interest (ROIs) were manually drawn using a custom MATLAB program by visually inspecting movies from all sessions and selecting neurons that showed at least one fluorescence transient in at least one session. Therefore, our analysis excludes neurons that do not show any fluorescence transients during our imaging periods. ROIs were aligned across sessions using a semi-automated method and classified as excitatory or inhibitory based on tdTomato expression. ROIs which presented with a nucleus filled by GCaMP5G at any point during the experiment, indicating possibly abnormal physiology<sup>39</sup>, were excluded from all analyses. Other than these nucleus-filled neurons, GCaMP-expressing neurons have been shown to display normal physiological properties including input resistance, resting membrane potential, input–output relationship, synaptic input maps, and normal synaptic plasticity<sup>8,39</sup>.

For structural imaging data, lateral motion for each image plane (20 frames) was corrected using full-frame cross-correlation image alignment (Turboreg<sup>38</sup> plugin for ImageJ), with the average of the five most consistent consecutive frames as the reference image. After this alignment, all 20 frames within a plane were averaged. Different image planes were then aligned using recursive alignment of stacks of images (Stackreg, plugin for ImageJ).



**Fluorescence analysis.** Pixels within each ROI were averaged to create a fluorescence time series. Background fluorescence fluctuations were subtracted from each ROI to remove neuropil contamination as follows. A ring-shaped 'background ROI' was created from the border of each neuronal ROI to a width of 6 pixels. From this background ROI, pixels containing transients that did not contaminate the neuronal ROI were excluded. These excluded pixels were identified as those that contained time points at which pixel values exceeded the neuronal ROI by two times the standard deviation of the difference between each background ROI pixel time series and the neuronal ROI time series. The remaining pixels were averaged to create a background fluorescence trace, and the  $\Delta F$  of the background fluorescence trace was subtracted from the neuronal ROI fluorescence trace to create the final background-subtracted fluorescence trace for each neuronal ROI.

The time-varying baseline ( $F_0$ ) of a fluorescence trace was estimated by smoothing inactive portions of the trace using the iterative procedure detailed below.

Inactive portions of the trace were initially estimated as when the raw trace loess-smoothed with a 1-s window was below a threshold. These inactive portions were further shortened by 5 s on each end to exclude tails of active portions. The threshold for activity was estimated by first creating a preliminary  $F_0$  approximation, which was a 1-min moving average of the original fluorescence trace. This preliminary  $F_0$  was subtracted from the raw trace to yield a preliminary  $\Delta F$ . The noise of the fluorescence was calculated as the standard deviation of the difference between the preliminary  $\Delta F$  and the smoothed preliminary  $\Delta F$  (loess, 1 s). An offset of the preliminary  $F_0$  was then estimated as the mode of the smoothed preliminary  $\Delta F$ . The threshold for detecting active portions was then set as preliminary  $F_0$  + the offset + two times the fluorescence noise. The remaining inactive portions of the trace were then concatenated and subjected to a second round of activity extraction using the same procedure, but by defining inactive portions where values fell within  $\pm$  two times the fluorescence noise.

Inactive portions were concatenated and smoothed (loess, 1 s). The resulting smoothed trace was then broken up according to their original time points and values were linearly interpolated across gaps (that is, active portions), resulting in an  $F_0$  estimation that was independent of activity and slow drifts. This  $F_0$  estimation was fine-tuned for an offset as follows. The  $F_0$  was subtracted from the raw trace, yielding a new  $\Delta F$ . The fluorescence noise was once again estimated by the standard deviation of the difference between  $\Delta F$  and smoothed  $\Delta F$  (loess, 1 s). The offset was then estimated by the mode of Gaussian fit of the distribution of values of inactive portions of the  $\Delta F$ . Inactive portions were defined as when the  $\Delta F$  values were within  $\pm$  two times the noise values, and were shortened by 5 s on either end. This offset was added to the  $F_0$  estimation, yielding the final  $F_0$ .

**Activity analysis.** Activity event traces were created from normalized background-subtracted fluorescence traces. For excitatory neurons, events were defined if the first derivative (velocity) of the smoothed fluorescence trace (loess, 1 s) crossed five times the standard deviation of the inactive velocity trace (inactive velocity trace was derived from periods when the fluorescence was within three times the standard deviation of the fluorescence trace). This velocity criterion detected sharp rises in the fluorescence trace. For these detected events, the start and end times were defined using the following iterative process. The peak time of the event was first roughly estimated as the time when the velocity drops below zero for the first time after it crossed the threshold as above. The peak time (that is, the end time of the event) was then defined as the time of the highest  $\Delta F/F_0$  value within five frames before and after the initial estimate. We next defined the 'baseline'  $\Delta F/F_0$  value for the event as the  $\Delta F/F_0$  value at the first time point when velocity was above zero before the peak time. (This 'baseline'  $\Delta F/F_0$  for each event is similar to the baseline of the fluorescence trace (zero) except in cases when the event occurs during a decay of another event; see the last event in Extended Data Fig. 4b for such an example.) The start time of the event was then defined as the last time point before the peak time when the  $\Delta F/F_0$  value is within noise level from the 'baseline'  $\Delta F/F_0$  ('noise' being three times the standard deviation of the difference between the raw  $\Delta F/F_0$  trace and the loess-smoothed  $\Delta F/F_0$  trace.) An activity event trace was then constructed which was zero except for frames with detected events, and each event was assigned an amplitude equal to the difference between the peak  $\Delta F/F_0$  and the 'baseline'  $\Delta F/F_0$  for that event. This eliminated the decay of the calcium signal<sup>19</sup>, but the use of velocity preserved events which occurred on top of the decays from other events. For inhibitory neurons, activity event traces were generated with the following procedure. The fluorescence noise was defined as the mean of the absolute difference between the raw trace and a 1 s moving window loess smoothed fluorescence trace. A low threshold of 1 times the noise and a high threshold of 3 times the noise were then set. Events were required to cross the high threshold. The start of an event was defined as the time when the fluorescence trace crossed the low threshold going up to capture the start of activity, and the end was defined as the time when the fluorescence trace crossed the high threshold going down. An activity event trace was then constructed which was zero at all frames except during detected events which were assigned values of the original  $\Delta F/F_0$  for those frames.

**Classification of movement-related neurons.** We observed higher levels of activity during movement periods. In individual neurons, the average percentages of image frames that contained activity during movement periods versus non-movement periods were  $0.68 \pm 0.03\%$  versus  $0.23 \pm 0.02\%$  in excitatory neurons and  $14.45 \pm 0.49\%$  versus  $6.95 \pm 0.27\%$  in inhibitory neurons (mean  $\pm$  s.e.m.). Neurons whose activity was significantly higher during movement periods were classified as movement-related using the following procedure. The amount of activity during movement was calculated for each neuron as the mean value of the activity event trace during movement epochs (defined as described above, and extending individual epochs by 5 image frames before and after each movement). The movement trace was then shuffled (10,000 times) such that complete movement epochs were kept intact but their position in the trace and relation to each other was randomized. A measure of activity during these shuffled movement epochs was calculated in each shuffle as above. The neuron was classified as movement-related if the real value was higher than the 0.5 percentile value of the shuffled values. Classification using only rewarded movements instead of all detected movements gave nearly identical results.

For analysing the longitudinal dynamics of the fractions of movement-related neurons over sessions (Fig. 2b), the fraction of movement-related neurons in each session in each animal was normalized to the highest and lowest values of the animal. The correlation was significantly positive for the data points in sessions 1–3 ( $r = 0.60$ ,  $P < 0.01$ ) and significantly negative for sessions 4–14 ( $r = -0.36$ ,  $P < 0.01$ ).

**Population activity correlation.** The stability of the population across days was assessed by correlation of population vectors (Fig. 2d). Each square represents the correlation coefficient of the excitatory neuron population activity vectors in a pair of sessions. The population activity vectors are the concatenation of the fraction of rewarded movements during which each neuron exhibited activity events. (For example, if neurons a, b and c are active in 20%, 70% and 30% of rewarded movements, respectively, then the population activity vector of the session is (0.2, 0.7, 0.3)).

**Activity onset timing analysis.** We performed two analyses to define the timing of activity of movement-related excitatory neurons relative to movement onset. To define when the population activity diverged from baseline, the activity of each neuron was first averaged across movements. Then the population activity vector in each image frame was compared to the baseline activity (all non-movement periods from all sessions) by bootstrap (10,000 repetitions), yielding a  $P$  value for each image frame. We identified image frames in which  $P$  values were below 0.01 in at least five successive frames, and defined the first frame of those as the time of divergence. The time of first population activity divergence from baseline was 105 ms before movement onset (Extended Data Fig. 6d). Similarly, to identify the timing of activity modulation of individual neurons, the activity of a neuron in each image frame across movements was compared to its baseline activity (420–315 ms before movement onset) by bootstrap (10,000 repetitions). We defined the activity onset timing as the first image frame in which  $P$  values were below 0.01 in at least five successive frames. 9.2% of movement-related excitatory neurons showed significant activity before movement onset (Extended Data Fig. 6f).

**Analysis of movement–activity correlation.** The relationship between movements and activity (Fig. 3) was analysed using 3 s (or 500 ms, Extended Data Fig. 7) starting from the onset of each rewarded movement, which was approximately the duration of rewarded movements for all animals over all sessions ( $2.62 \pm 0.02$  s, mean  $\pm$  s.e.m.). The learned movement and activity patterns were created by averaging the lever traces and the activity of excitatory neurons, respectively, of a randomly chosen half of the trials of sessions 10–14, considered the expert stage. The trials used to generate the learned patterns were excluded from correlation analysis. The movement correlation for each trial was the correlation coefficient of the lever trace of the rewarded movement in that trial with the learned movement pattern. The activity correlation for each trial was the correlation coefficient between the concatenated activity time series of all excitatory neurons in the trial and the concatenated learned activity pattern. For results shown in Fig. 3b, the random choice of trials to define the learned patterns was repeated 1,000 times and the results were averaged.

**Dendritic spine dynamics.** Dendritic spines were manually scored over the entire 14 training sessions using a custom-written IGOR program (J. Boyd and K. Haas). Spine analysis was done in three dimensions and the criteria were as previously described<sup>21</sup>. Analysis was done blind to the session number of each image, which was randomized. We assumed that rapid 'flickering' of spines (elimination and immediate reformation, or formation and immediate elimination) is rare and corrected our blind scoring accordingly. While this corrected for mistakes in scoring, we may be slightly underestimating spine dynamics. Specifically, if a spine was scored as absent in one session (session X) and present in the immediately preceding (session X–1) and following (session X+1) sessions, then it was called present on session X. Furthermore, if a spine was scored as present in one session (session X) and absent in the immediately preceding (session X–1) and following (session X+1) sessions, then it was called absent on session X. No more than one correction was applied on any given spine. If a spine score contained these gaps after one correction, that spine was excluded from following analyses. These exclusions were rare (4 of 191). 8% of spines were

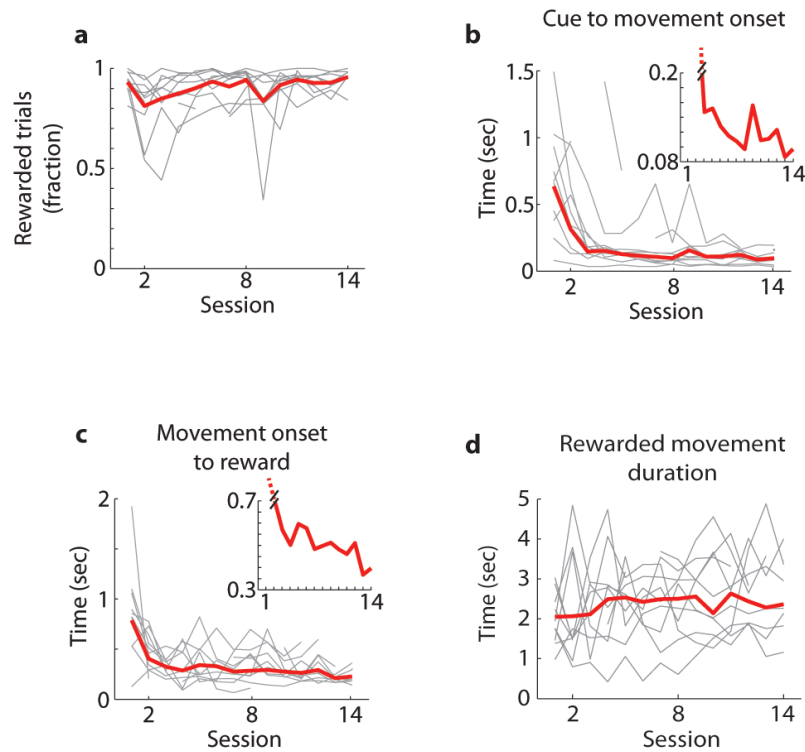
corrected (16 of 191). The results closely matched those from independent scoring of the same data without shuffling dates (data not shown).

**Simultaneous two-photon guided cell-attached recordings and calcium imaging.** Mice that had previously been used for calcium imaging were allowed full access to water before electrophysiology. On the day of the experiment, mice were anaesthetized with isoflurane and the glass window was removed and replaced with a glass half-window which was secured with superglue. The animals were then head-fixed in the imaging rig and allowed to recover from anaesthesia. Loose patch recordings were performed with glass pipettes ( $\sim 5\text{--}7\text{ M}\Omega$ ) filled with  $100\text{ }\mu\text{M}$  Alexa Fluor 488 in saline. Excitatory neurons (negative for tdTomato) expressing GCaMP5G without fluorescence in the nucleus were targeted for recording. Signals were amplified  $500\times$  by an Axon CNS amplifier (Molecular Devices), filtered at  $2\text{ kHz}$ , recorded (Ephus) at  $10\text{ kHz}$ , and synchronized to the start of image acquisition. In 4 out of 6 neurons, imaging was done at the same zoom as the population imaging experiments (field of view  $472 \times 508\text{ }\mu\text{m}$ ), and at three times higher zoom for the other two neurons. The results were similar at both zooms.

**Statistics.** Non-parametric tests were used when possible to avoid assumptions about data distributions. Sample sizes were determined based on the statistical significance of our main findings, which is highly significant. Multiple comparisons were corrected for by Bonferroni corrections. Sample sizes ( $n$ ) are as follows where applicable. Mice per session: 6, 6, 7, 7, 7, 7, 7, 7, 7, 6, 6, 5. Imaged, rewarded trials without movement at cue/total trials per session: 98/313, 173/526, 458/968, 488/1,071, 438/966, 519/1,081, 424/946, 453/1,082, 320/784, 457/981, 481/889, 417/959, 412/853, 280/702. Movement-related/all imaged excitatory neurons: 79/874, 118/874,

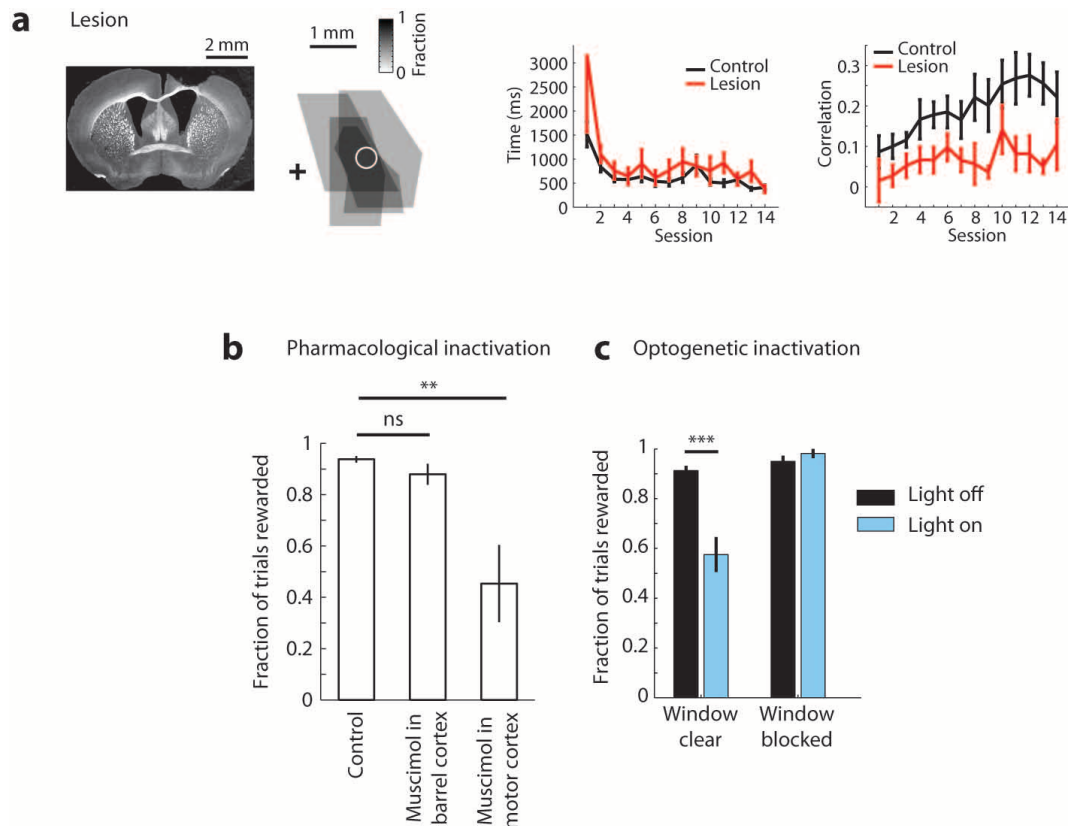
202/1,122, 210/1,122, 201/1,122, 213/1,122, 181/1,122, 191/1,122, 159/1,122, 171/1,122, 158/1,122, 136/995, 149/995, 89/843. Movement-related/all imaged inhibitory neurons: 103/262, 117/262, 155/297, 167/297, 154/297, 133/297, 133/297, 124/297, 106/297, 125/297, 119/297, 119/260, 114/260, 68/183.

31. Holtmaat, A. *et al.* Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. *Nature Protocols* **4**, 1128–1144 (2009).
32. Li, C. X. & Waters, R. S. Organization of the mouse motor cortex studied by retrograde tracing and intracortical microstimulation (ICMS) mapping. *Can. J. Neurol. Sci.* **18**, 28–38 (1991).
33. Pronichev, I. V. & Lenkov, D. N. Functional mapping of the motor cortex of the white mouse by a microstimulation method. *Neurosci. Behav. Physiol.* **28**, 80–85 (1998).
34. Ferezou, I. *et al.* Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. *Neuron* **56**, 907–923 (2007).
35. Ayling, O. G., Harrison, T. C., Boyd, J. D., Goroshkov, A. & Murphy, T. H. Automated light-based mapping of motor cortex by photoactivation of channelrhodopsin-2 transgenic mice. *Nature Methods* **6**, 219–224 (2009).
36. Tennant, K. A. *et al.* The organization of the forelimb representation of the C57BL/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture. *Cereb. Cortex* **21**, 865–876 (2011).
37. Olsen, S. R., Bortone, D. S., Adesnik, H. & Scanziani, M. Gain control by layer six in cortical circuits of vision. *Nature* **483**, 47–52 (2012).
38. Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
39. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).



**Extended Data Figure 1 | Behaviour.** The fraction of rewarded trials is consistently high but the timing of behaviour improves during learning. **a**, Fraction of trials that are rewarded. **b**, Time from cue onset to movement onset decreases ( $P < 0.001$ , one-way ANOVA); inset, zoom. **c**, Time from

movement onset to reward decreases ( $P < 0.001$ , one-way ANOVA); inset, zoom. **d**, The duration of each rewarded movement is stable throughout learning ( $P = 0.94$ , one-way ANOVA). Grey, individual mice; red, mean of all animals (**a**) or median of all trials (**b–d**).

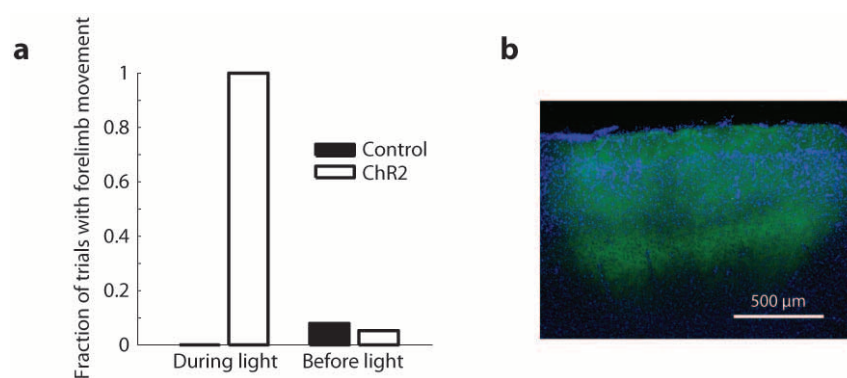


### Extended Data Figure 2 | Motor cortex is required for the lever-press task.

**a**, Aspiration lesion of motor cortex impairs learning. Mice were allowed to recover for 14 days after lesion before training. Left: histological image showing lesion in the right motor cortex and quantification of lesion extents in four mice shown as a density map of the fraction of animals in which the area was lesioned. Anterior is to the top; lateral to the right. + denotes bregma. The white circle indicates the imaged area. Middle: average time from movement onset to reward throughout learning. This time is longer in mice with motor cortex lesion ( $P < 0.01$ , two-way ANOVA), indicating that the mice with a lesion are less efficient in their movements. Right: correlation of lever movements in all pairs of trials within each session throughout learning. This correlation is lower in the mice with a lesion ( $P < 0.001$ , two-way ANOVA), indicating that the mice with a lesion do not develop reproducible movements. **b**, Injections of

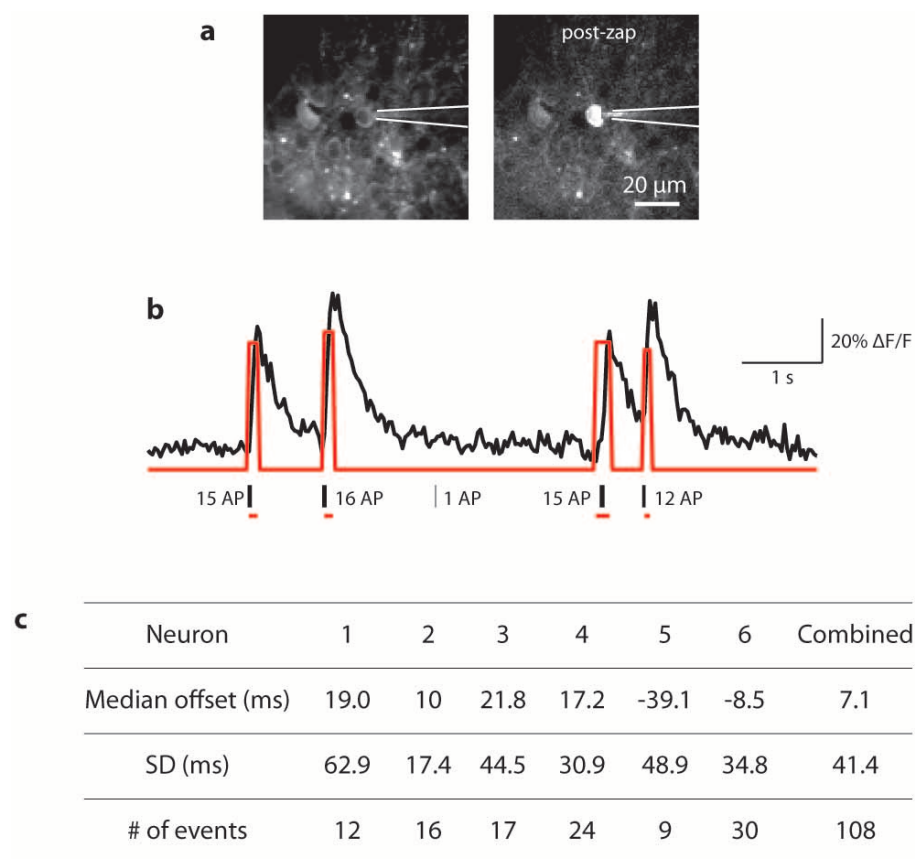
muscimol, a GABA receptor agonist, into the imaged area acutely impairs performance (control versus muscimol in motor cortex,  $**P < 0.01$ , Wilcoxon rank sum test). Muscimol injections in the barrel cortex had no significant effect (control versus muscimol in barrel cortex,  $P = 0.35$ , Wilcoxon rank sum test). Control,  $n = 18$  sessions in 6 mice; barrel cortex,  $n = 6$  sessions in 6 mice; motor cortex,  $n = 6$  sessions in 6 mice. **c**, The imaged cortical area was acutely inactivated by stimulation of ChR2 in parvalbumin-expressing inhibitory neurons by blue light in interleaved 20% of trials ( $n = 10$  sessions in 2 animals). This optogenetic inactivation of the imaged area impaired performance on a trial-by-trial basis ( $***P < 0.001$ , Wilcoxon rank sum test). Blue light had no effect on behaviour when the window was covered with opaque silicone ( $n = 4$  sessions in 2 animals). All error bars are s.e.m.





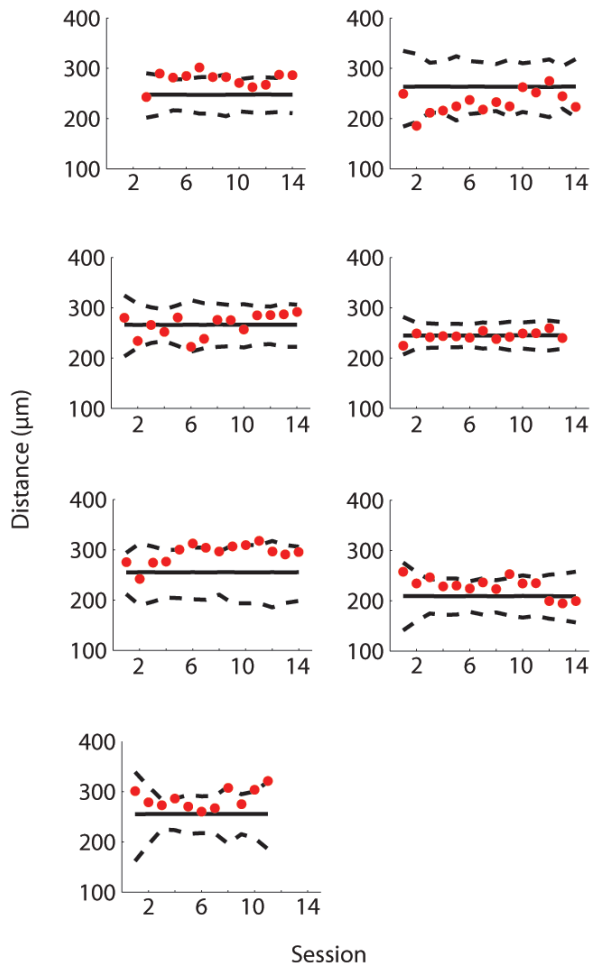
**Extended Data Figure 3 | Optogenetic stimulation of the imaged area evokes forelimb movements in awake mice.** **a**, Optogenetic excitation of the imaged area triggers forelimb movements in mice expressing ChR2 but not in control mice not expressing ChR2 ( $P < 0.001$ , chi-squared test). ChR2 expression does not alter spontaneous movement frequency in the absence of

stimulation ( $P = 0.64$ , chi-squared test). 'During light', 1-s light stimulation; 'before light', 2 to 1 s before light onset,  $n = 40$  'during light' trials and 38 'before light' trials in two ChR2 mice, 38 'during light' trials and 38 'before light' trials in two control mice. **b**, Histological section showing the expression of ChR2 in the motor cortex. Green, ChR2-YFP; blue, DAPI.

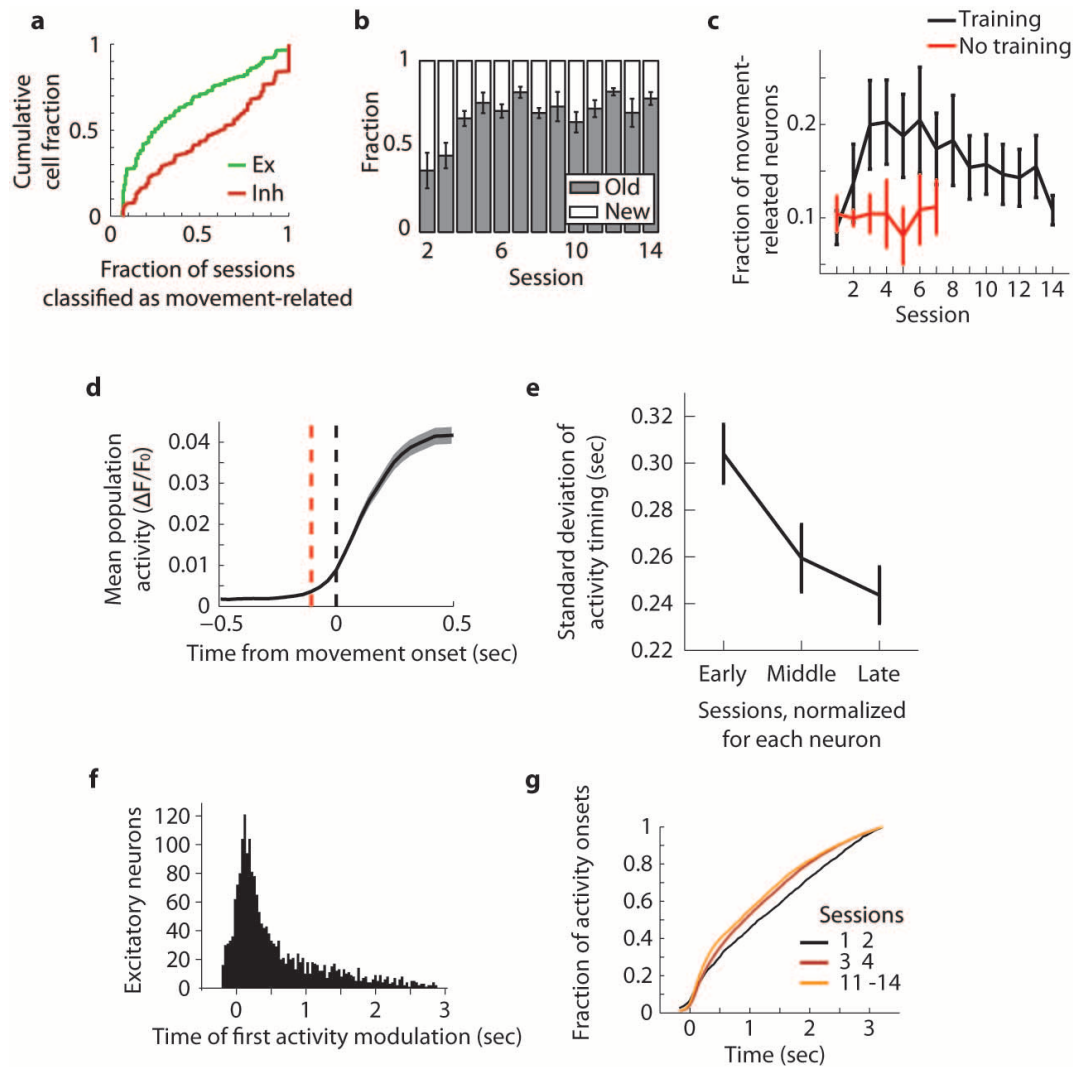


**Extended Data Figure 4 | Simultaneous cell-attached recordings and two-photon calcium imaging in awake mice.** **a**, Left: *in vivo* two-photon image of motor cortex neurons expressing GCaMP5G. The neuron in the centre is targeted with a patch electrode. Right: after the recording session, voltage step was applied to the electrode to activate the recorded neuron. The increased GCaMP5G fluorescence in the middle neuron confirms that the neuron was indeed targeted. **b**, Example GCaMP5G fluorescence trace (top: black indicates fluorescence trace and red indicates detected calcium events) and

simultaneously recorded action potentials (bottom: black vertical ticks; the numbers indicate the number of action potentials contained in each burst). Horizontal red lines at bottom indicate the duration of detected calcium events. Note the precise temporal relationship between action potentials and calcium events. **c**, Table summarizing data from six neurons in two mice. Positive offsets indicate the lag of the onset of detected calcium events relative to the first spike in the burst. The offset ( $7.1 \pm 41.4$  ms) is on the order of the temporal resolution of our imaging ( $\sim 35$  ms per image frame).



**Extended Data Figure 5 | Lack of spatial clustering of movement-related excitatory neurons.** Each plot represents one animal. Red dots show the mean pairwise distance between movement-related excitatory neurons. Solid and dotted black lines show the mean and 95% confidence intervals, respectively, obtained from shuffling the identities of movement-related neurons among all excitatory neurons 10,000 times. Dots below the lower dotted line would indicate significant clustering of cells, whereas dots above the upper dotted line would indicate the significant dispersion of cells ( $P < 0.05$ ).

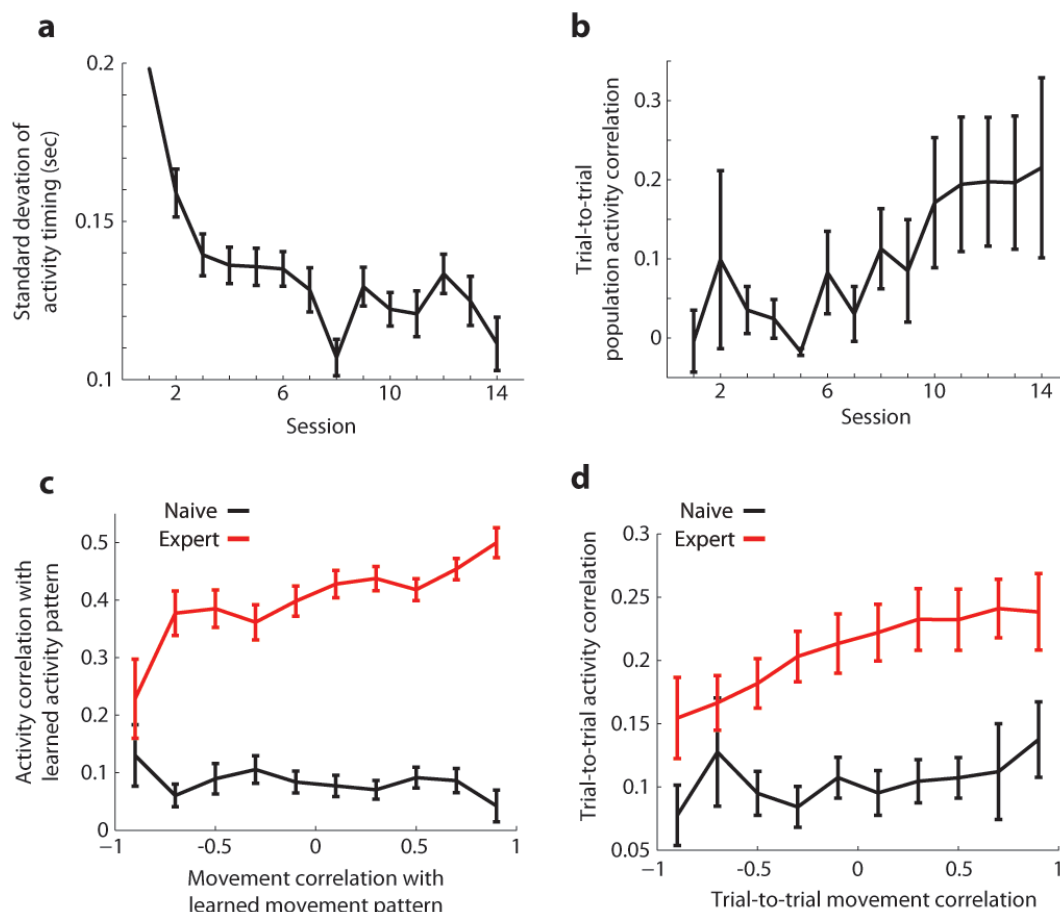


#### Extended Data Figure 6 | Additional analysis of population activity.

**a**, Cumulative distribution of fraction of sessions classified as movement-related for inhibitory (red) and excitatory (green) neurons, showing the relative invariance of inhibitory neurons and dynamism of excitatory neurons ( $P < 0.001$ , Kolmogorov–Smirnov test). **b**, Movement-related excitatory neuron populations in each session compared to the previous session. Grey, fraction of neurons classified in the previous session; white, not classified in the previous session. A large number of newly movement-related neurons were added in the first few sessions ( $P < 0.001$ , comparison between sessions 2–4 versus 10–14, Wilcoxon rank sum test). **c**, Fraction of excitatory neurons classified as movement-related in each session. Black, training ( $n = 7$  mice, this is the data shown in Fig. 2b); red, no training ( $n = 6$  mice). The expansion of movement-related neurons is specific to animals that underwent training ( $P = 0.74$ , sessions 1–2 combined;  $P < 0.001$ , sessions 3–7 combined; Wilcoxon rank sum test). **d**, Average population activity aligned to movement onset (black dotted line). Average activity (calcium event trace) of each movement-related excitatory neuron was averaged. The population activity diverged from baseline 105 ms before movement onset (red dotted line,

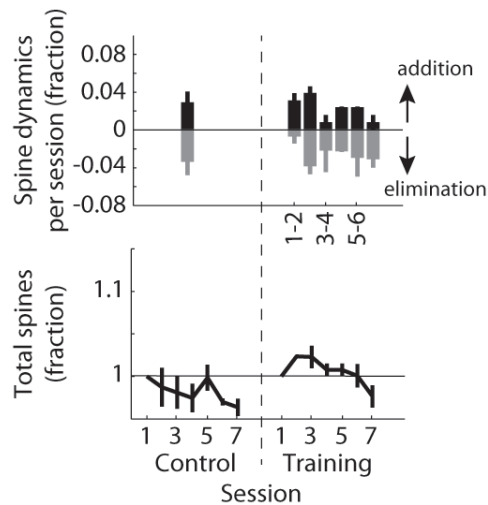
Methods). **e**, Standard deviation of activity timing of individual movement-related excitatory neurons across sessions. Focusing on neurons that are classified as movement-related in three or more sessions, the standard deviation of activity onset timing relative to movement onset is plotted across sessions. Sessions were binned into one-third of the total number of sessions each neuron was classified. Activity timing became more stable on the neuron-by-neuron basis ( $r = -0.14$ ,  $P < 0.001$ ). **f**, Histogram of the time from movement onset that the activity of each movement-related neuron significantly diverged from baseline. 9.2% of movement-related excitatory neurons show significant pre-movement activity, a composition similar to a previous study<sup>14</sup>. 82.7% of activity of movement-related neurons occurred during the periods between 105 ms before movement onset and movement offset (Methods). **g**, The cumulative fraction plot of the timing of all activity onsets of movement-related excitatory neurons during rewarded movements. Each group of sessions is shown as a line, with different colours representing different sessions. The distribution of activity onset timing during later sessions shifts towards the movement onset ( $P < 0.001$ , Kolmogorov–Smirnov test for all three comparisons). All error bars are s.e.m.



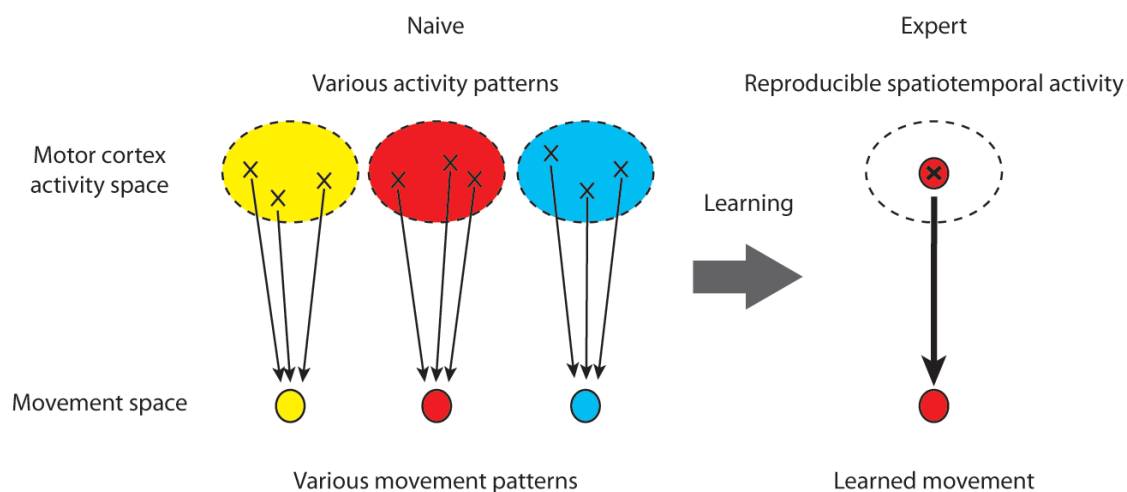


**Extended Data Figure 7 | Activity analysis focusing on the first 500 ms of each movement.** For the activity analyses in the main figures that used the duration of 3 s after movement onset, we repeated the same analyses focusing on the first 500 ms of each movement (median time from movement onset to reward = 506 ms). This early activity shows progression throughout learning, similar to when activity over 3 s was considered. **a**, Standard deviation of the timing of activity onsets for movement-related excitatory neurons over sessions, indicating a gradual refinement of activity timing ( $r = -0.18$ ,  $P < 0.001$ ). Neurons that were active in less than five trials of a given session were excluded from this analysis. The first bin contains only one data point and thus does not have an error bar. This analysis is equivalent to Fig. 2g. **b**, Pairwise trial-to-trial correlation of temporal population activity vectors increases with learning ( $r = 0.38$ ,  $P < 0.001$ ). Temporal population activity vector was defined as a concatenation of the activity traces of all movement-related

neurons and thus maintained temporal information within each movement. This analysis is equivalent to Fig. 2h. **c**, Correlation of spatiotemporal activity with the learned activity pattern is a function of the correlation of movement with the learned movement pattern in expert sessions. Movements similar to the learned movement pattern but made in naive sessions display activity very different from the learned activity pattern ( $P = 0.28$  and  $< 0.001$  in the bins 1 and 2–10, respectively, Wilcoxon rank sum test). This analysis is equivalent to Fig. 3b. **d**, Pairwise trial-to-trial correlation of temporal population activity vectors plotted as a function of movement correlation on those trials. A strong relationship between population activity and movement emerges during learning ( $P = 0.08$ ,  $= 0.08$ ,  $= 0.004$ ,  $< 0.001$ ,  $= 0.002$ ,  $< 0.001$ ,  $= 0.001$ ,  $= 0.002$ ,  $< 0.001$  and  $= 0.046$  for each bin, Wilcoxon rank sum test). This analysis is equivalent to Fig. 3c. All error bars are s.e.m.



**Extended Data Figure 8 | Dynamics of dendritic spines in the hindlimb area during learning of the lever-press task.** Summary of dendritic spine dynamics in the hindlimb area during control period (7 days before training) and subsequent 7 days of training. Mice were water restricted in both conditions. Top: spine additions (black) and eliminations (grey) in each session. For control sessions, data from all sessions are combined. Bottom: total spine number across sessions. Values are normalized to the total spine number in session 1 in each condition. Unlike the forelimb area, the density of dendritic spines in the hindlimb area is relatively stable during learning ( $P = 0.07$ , comparisons between control versus training sessions 4–7, Wilcoxon rank sum test). All error bars are s.e.m.



**Extended Data Figure 9 | Schematic of learning-related changes in the relationship of motor cortex activity and movement.** Top: abstract space of activity patterns. Bottom: abstract space of movements. Circles in the movement space represent observed movements, and ovals in the activity space represent possible activity patterns that can lead to corresponding movements. Crosses and arrows represent example individual trials of activity–movement pairs. In naive animals, each trial involves variable activity and movement

patterns as illustrated by scattered crosses and multiple movements. In this stage, the relationship between activity and movement is inconsistent (that is, degenerate), such that same movement is derived from different activity patterns in different trials. During learning, this degeneracy is reduced and a reproducible spatiotemporal activity pattern emerges in the motor cortex that reliably generates the learned movement. This learned activity pattern (bold cross) is rarely, if at all, observed in naive stages.

# The unfolded protein response governs integrity of the haematopoietic stem–cell pool during stress

Peter van Galen<sup>1,2</sup>, Antonija Kreso<sup>1,2</sup>, Nathan Mbong<sup>1,2</sup>, David G. Kent<sup>3</sup>, Timothy Fitzmaurice<sup>4</sup>, Joseph E. Chambers<sup>5</sup>, Stephanie Xie<sup>1,2</sup>, Elisa Laurenti<sup>1,2</sup>, Karin Hermans<sup>1,2</sup>, Kolja Eppert<sup>6</sup>, Stefan J. Marciniak<sup>5</sup>, Jane C. Goodall<sup>4</sup>, Anthony R. Green<sup>3</sup>, Bradley G. Wouters<sup>7</sup>, Erno Wienholds<sup>1,2</sup> & John E. Dick<sup>1,2</sup>

The blood system is sustained by a pool of haematopoietic stem cells (HSCs) that are long-lived due to their capacity for self-renewal. A consequence of longevity is exposure to stress stimuli including reactive oxygen species (ROS), nutrient fluctuation and DNA damage<sup>1,2</sup>. Damage that occurs within stressed HSCs must be tightly controlled to prevent either loss of function or the clonal persistence of oncogenic mutations that increase the risk of leukaemogenesis<sup>3,4</sup>. Despite the importance of maintaining cell integrity throughout life, how the HSC pool achieves this and how individual HSCs respond to stress remain poorly understood. Many sources of stress cause misfolded protein accumulation in the endoplasmic reticulum (ER), and subsequent activation of the unfolded protein response (UPR) enables the cell to either resolve stress or initiate apoptosis<sup>5,6</sup>. Here we show that human HSCs are predisposed to apoptosis through strong activation of the PERK branch of the UPR after ER stress, whereas closely related progenitors exhibit an adaptive response leading to their survival. Enhanced ER protein folding by overexpression of the co-chaperone ERDJ4 (also called *DNAJB9*) increases HSC repopulation capacity in xenograft assays, linking the UPR to HSC function. Because the UPR is a focal point where different sources of stress converge, our study provides a framework for understanding how stress signalling is coordinated within tissue hierarchies and integrated with stemness. Broadly, these findings reveal that the HSC pool maintains clonal integrity by clearance of individual HSCs after stress to prevent propagation of damaged stem cells.

The human haematopoietic hierarchy has recently been delineated at the single-cell level, enabling precise isolation of HSCs and progenitor cells<sup>7–9</sup> (Extended Data Table 1). Pathway analysis using gene expression signatures from lineage-depleted umbilical cord blood populations revealed enrichment of UPR components in HSCs compared to progenitor cells<sup>10</sup> (Fig. 1a and Extended Data Fig. 1a). The UPR encompasses the IRE1, PERK and ATF6 pathways<sup>6</sup> (Extended Data Fig. 1b). Several genes of the PERK signalling branch were more highly expressed in a mixed population of HSCs and progenitor cells (CD34<sup>+</sup>CD38<sup>–</sup> HSPCs) compared to downstream progenitors (CD34<sup>+</sup>CD38<sup>+</sup>) isolated from cord blood and adult bone marrow (Fig. 1b and Extended Data Fig. 1c–g). However, splicing of *XBPI* messenger RNA, which is representative of IRE1 activity, was lower in HSPCs compared to progenitors (spliced/total *XBPI*, Fig. 1b). Taken together, gene expression analysis of HSPC and progenitor fractions indicates differential activation of UPR branches, with increased expression of PERK-dependent genes and decreased activity of IRE1 in HSPCs.

To examine whether differential basal UPR gene expression reflects distinct ER stress responses in HSPCs and progenitors, we used two chemical inducers of ER stress: thapsigargin and tunicamycin. Thapsigargin disrupts Ca<sup>2+</sup> homeostasis in the ER, rapidly activating all three branches of the UPR<sup>11</sup>. Treatment of sorted HSPCs and progenitors with

thapsigargin resulted in upregulation of the canonical UPR target genes *GRP94* (also called *HSP90B1*), *GRP78* (*HSPA5*) and *ERDJ4* (Fig. 1c and Extended Data Fig. 2a). *XBPI* mRNA was rapidly spliced in progenitors but to a lesser extent in HSPCs. This indicates that differential *XBPI* splicing between HSPCs and progenitors under steady-state conditions is exaggerated upon thapsigargin-induced ER stress, consistent with repressed IRE1 splicing activity in HSPCs.

Tunicamycin blocks synthesis of N-linked glycoproteins, causing accumulation of unfolded proteins in the ER<sup>11</sup>. Tunicamycin treatment resulted in higher upregulation of the canonical UPR genes *GRP94*, *GRP78* and *ERDJ4* in HSPCs compared to progenitors (Fig. 1d and Extended Data Fig. 2b). Furthermore, upregulation of the PERK pathway constituents *CHOP* (also called *DDIT3*), *ATF4* and *GADD34* (*PPP1R15A*) was higher in HSPCs compared to progenitors. In adult bone marrow, *CHOP* expression was also higher in HSPCs compared to progenitors after addition of a high tunicamycin dose (Extended Data Fig. 2c). Thus, basal enrichment of PERK pathway target genes in HSPCs is further amplified with tunicamycin treatment.

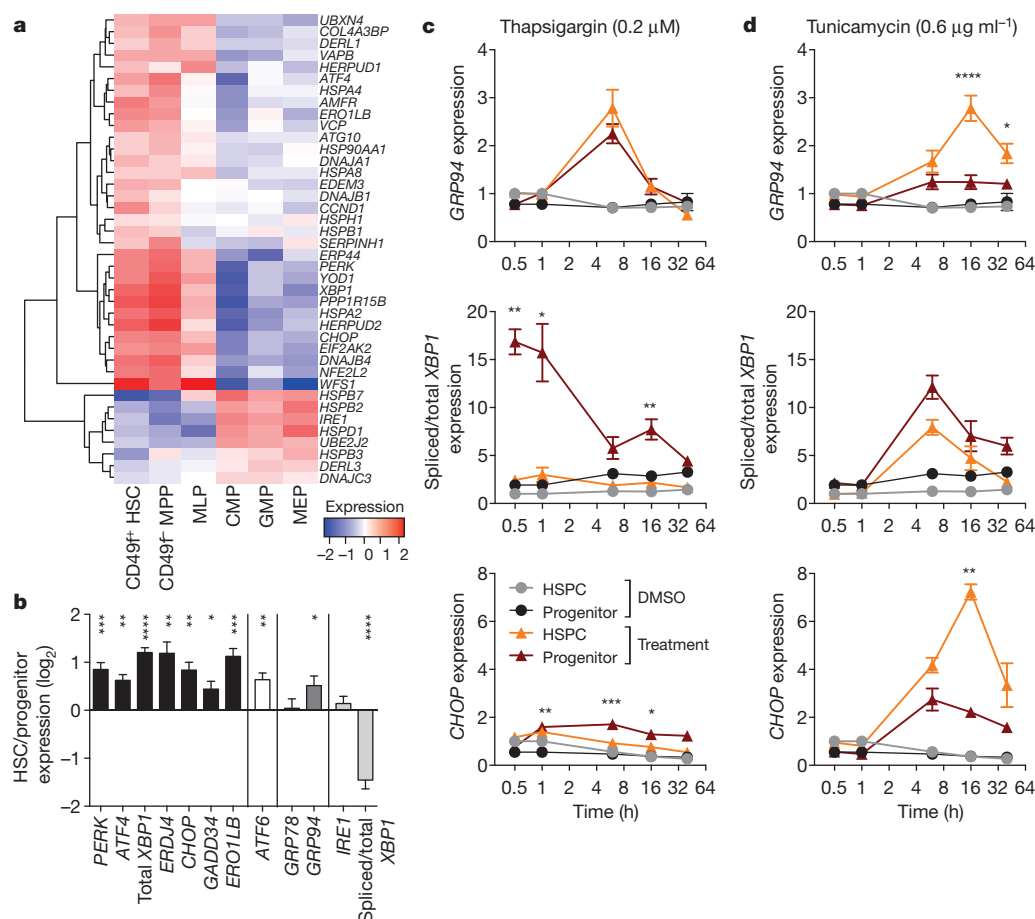
Because persistent ER stress can lead to activation of apoptosis through signals downstream of the IRE1 and PERK branches of the UPR<sup>12</sup>, differential UPR branch activation between HSPCs and progenitors might influence cell fate outcome. Thapsigargin treatment did not result in survival differences between HSPCs and progenitors (Extended Data Fig. 3a). However, tunicamycin treatment significantly reduced HSPC viability and clonogenic capacity as compared to progenitors (Fig. 2a–c and Extended Data Fig. 3b, c). Activation of HSPCs into the cell cycle with cytokines reduced the survival difference with progenitors, indicating that tunicamycin sensitivity may be linked to the inherent quiescence of HSPCs (Extended Data Fig. 3d). By focusing on highly purified fractions, we found that tunicamycin caused selective loss by apoptosis of phenotypic CD34<sup>+</sup>CD38<sup>–</sup>CD45RA<sup>–</sup>CD90<sup>+</sup> HSCs (Fig. 2d and Extended Data Fig. 3e, f). Overall, these data indicate that tunicamycin-induced ER stress not only elicits distinct UPR signalling in HSCs compared to progenitors, but also causes selective apoptosis of HSCs.

ER stress induces eIF2 $\alpha$  phosphorylation (p-eIF2 $\alpha$ ) by PERK, leading to global translational attenuation but, paradoxically, ATF4 and CHOP translation is increased<sup>13–15</sup>. ATF4 and CHOP can induce apoptosis following prolonged ER stress, in part by upregulating the eIF2 $\alpha$  phosphatase *GADD34*, leading to increased protein load through translational recovery<sup>16,17</sup> (Extended Data Fig. 1b). We investigated whether the increased apoptosis of HSCs compared to progenitors is linked to preferential PERK pathway activation. A lentiviral reporter vector was constructed to measure the ATF4 translation rate<sup>13</sup> (Fig. 3a and Extended Data Fig. 4a–c). As expected, the increased ATF4 reporter activation that occurs after tunicamycin treatment was inhibited by PERK inhibition (measured by transgene ratio, Extended Data Fig. 4d). This activation was more efficient in HSPCs compared to progenitors (Fig. 3b, c),

<sup>1</sup>Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario M5G 2M9, Canada. <sup>2</sup>Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

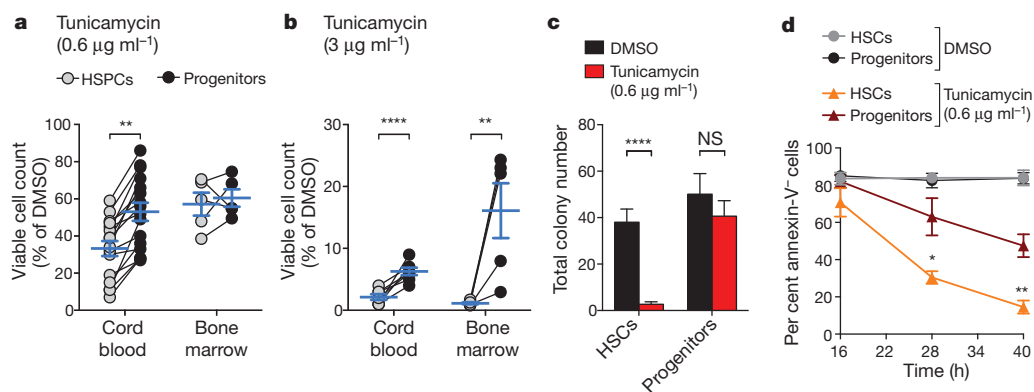
<sup>3</sup>Cambridge Institute for Medical Research, Wellcome Trust/MRC Stem Cell Institute and Department of Haematology, University of Cambridge, Cambridge CB2 0XY, UK. <sup>4</sup>Department of Medicine, School of Clinical Medicine, Addenbrookes Hospital, University of Cambridge, Cambridge CB2 0QQ, UK. <sup>5</sup>Cambridge Institute for Medical Research, Wellcome Trust/MRC Stem Cell Institute and Department of Medicine, University of Cambridge, Cambridge CB2 0XY, UK. <sup>6</sup>Department of Pediatrics, McGill University and the Research Institute of the McGill University Health Centre, Westmount, Québec H3Z 2Z3, Canada. <sup>7</sup>Departments of Radiation Oncology and Medical Biophysics, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario M5G 2M9, Canada.





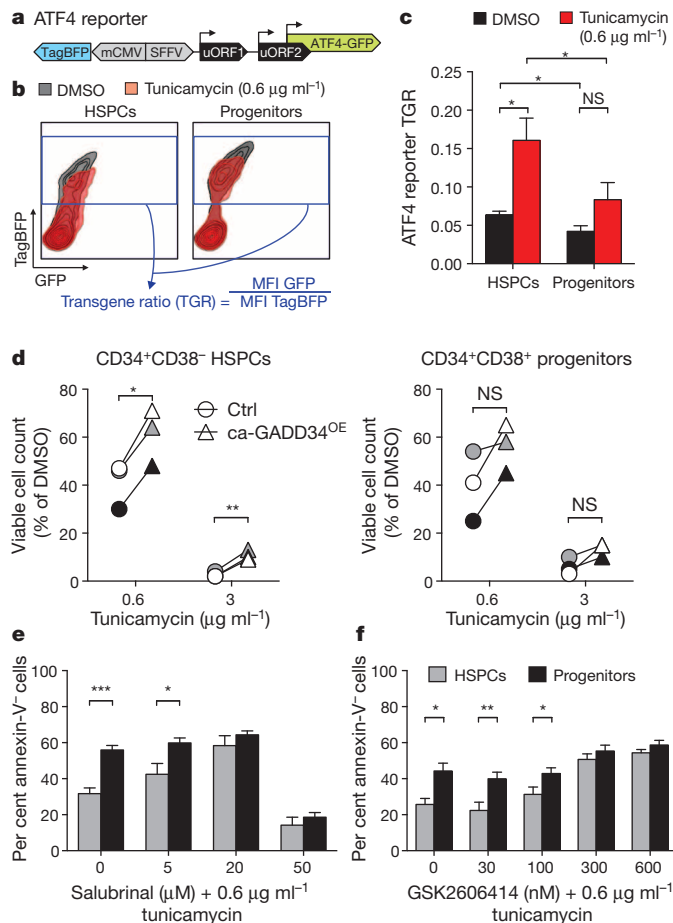
**Figure 1 | Elevated expression of PERK branch genes of the UPR in HSCs compared to progenitors and further amplification after tunicamycin-induced stress.** **a**, Forty UPR-related genes from the nodes in Extended Data Fig. 1a showed differential expression between HSCs and progenitors (false discovery rate (FDR) <0.05). CMP, common myeloid progenitor; GMP, granulocyte macrophage progenitor; MEP, megakaryocyte erythrocyte progenitor; MPP, multipotent progenitor; MLP, multilymphoid progenitor. **b**, Expression of key UPR genes in HSPC and progenitor fractions was measured by qPCR. Results are shown as mean  $\pm$  s.e.m. of  $n = 6$  cord

blood samples. **c**, **d**, UPR branch activation depends on cell type and stressor. Sorted HSPCs or progenitors were plated in the presence of thapsigargin (**c**) or tunicamycin (**d**). RNA was isolated at different time points to measure gene expression by qPCR. DMSO controls were the same between **c** and **d**. Data are shown as mean  $\pm$  s.e.m. of  $n = 3$  cord blood samples;  $P$  value was calculated based on treated/control cells and indicates differential response between HSPCs and progenitors. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .



**Figure 2 | HSCs are predisposed to apoptosis compared to progenitors after treatment with the ER stress agent tunicamycin.** **a**, **b**, Lower survival of cord-blood- and bone-marrow-derived HSPCs compared to progenitors in the presence of tunicamycin. Sorted HSCs/HSPCs and progenitors were plated with 0.6  $\mu$ g ml<sup>-1</sup> tunicamycin (**a**) or 3  $\mu$ g ml<sup>-1</sup> tunicamycin (**b**). Symbols represent viable cell counts of individual samples where populations are connected by a black line; the blue line indicates mean  $\pm$  s.e.m. of  $n = 16$  cord blood samples and  $n = 5$  bone marrow samples (**a**) or  $n = 7$  cord blood samples and  $n = 5$  bone marrow samples (**b**). **c**, Reduced clonogenic potential

of HSCs compared to progenitors following tunicamycin treatment. HSCs or progenitors were sorted into methylcellulose containing DMSO or tunicamycin. Data are shown as mean  $\pm$  s.e.m. of  $n = 4$  cord blood samples; NS, not significant. **d**, Tunicamycin treatment causes higher apoptosis in HSCs compared to progenitors. Cord blood cells were plated with tunicamycin and stained for primitive surface markers and Annexin-V/Sytox. Quantification of viable cells is shown as mean  $\pm$  s.e.m. of  $n = 5$  cord blood samples.  $P$  values indicate different viability between HSCs and progenitors. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ .



**Figure 3 | HSCs are predisposed to UPR-induced apoptosis through PERK-eIF2 $\alpha$ -ATF4-CHOP-GADD34 signalling.** **a**, Bidirectional lentiviral reporter vector. All transduced cells are marked by TagBFP; GFP brightness measures the ATF4 mRNA translation rate, which is regulated by upstream open reading frames (uORFs) and depends on pElF2 $\alpha$  (ref. 13). **b**, **c**, Higher ATF4 reporter activity in HSPCs compared to progenitors. HSPCs and progenitors were sorted, transduced with the ATF4 reporter and treated with tunicamycin. **b**, Representative flow plots outline calculation of the transgene ratio. MFI, mean fluorescence intensity. **c**, Results are shown as mean  $\pm$  s.e.m. of  $n = 4$  cord blood samples. **d**, Constitutively active GADD34<sup>OE</sup> has a more pronounced effect on HSPCs compared to progenitors. Transduced HSPCs (left) or progenitors (right) were treated with tunicamycin. Symbols represent  $n = 3$  cord blood samples where control (ctrl) and constitutively active GADD34<sup>OE</sup> (ca-GADD34<sup>OE</sup>) groups are connected by a black line;  $P$  values were calculated using paired  $t$ -tests. **e**, **f**, Modulating the PERK pathway rescues HSCs from apoptosis. HSPCs and progenitors were plated with tunicamycin and the GADD34 inhibitor salubrinal (**e**) or the PERK inhibitor GSK2606414 (**f**). Viability is shown as mean  $\pm$  s.e.m. of  $n = 5$  (**e**) or  $n = 4$  (**f**) cord blood samples. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

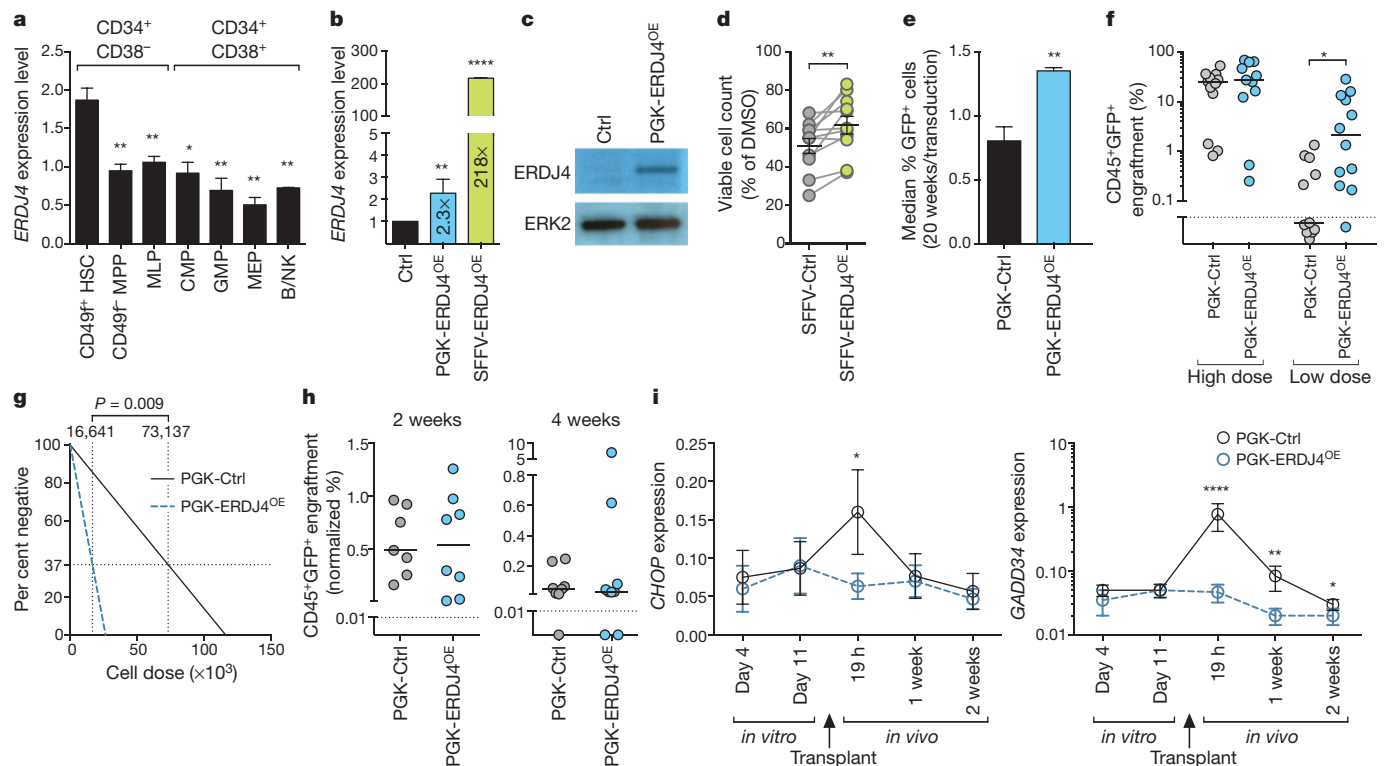
indicating that tunicamycin stimulates PERK pathway activity more strongly in HSPCs. As a second method to test PERK pathway involvement, we overexpressed constitutively active GADD34, which prevents pElF2 $\alpha$  and upregulation of ATF4 and CHOP<sup>18</sup>. Expression of constitutively active GADD34 significantly increased survival of HSPCs, but not progenitors, after tunicamycin treatment (Fig. 3d). In a third independent approach, HSPCs and progenitors were treated with both tunicamycin and salubrinal, which prevents eIF2 $\alpha$  dephosphorylation<sup>19</sup>. Addition of salubrinal preferentially increased HSPC survival after tunicamycin treatment (Fig. 3e). Finally, HSPCs and progenitors were treated with both tunicamycin and the PERK inhibitor GSK2606414 (ref. 20). Like salubrinal, GSK2606414 reduced the survival difference between HSPCs and progenitors (Fig. 3f). Thus, interfering with the PERK pathway at multiple junctions protects HSPCs from tunicamycin-induced

apoptosis and equalizes the survival between HSPCs and progenitors. Collectively, these data demonstrate that ER stress preferentially induces apoptosis of HSPCs compared to closely related progenitors through selective activation of the PERK branch of the UPR.

We next asked whether the UPR was involved in regulating HSC function *in vivo*. CHOP is a main driver of apoptosis following PERK activation<sup>12,16</sup> and analysis of *Chop*<sup>-/-</sup> mouse bone marrow showed a small increase in the viability of mouse HSPCs (Extended Data Fig. 5a, b and Extended Data Table 2). This result indicates that *Chop* may be required for the survival/death balance of mouse HSPCs under physiological conditions. Next, we investigated whether enhanced ER protein folding would alter human HSC function. ERDJ4 increases the activity of the chaperone GRP78 and can associate with the ERAD machinery<sup>21,22</sup>. These functions may enhance ER protein folding capacity and protect against UPR-induced apoptosis<sup>23</sup>. ERDJ4 expression was highest in CD49f<sup>+</sup> HSCs and reduced in all downstream progenitors (Fig. 4a). Green fluorescent protein (GFP)-marked ERDJ4 overexpression (ERDJ4<sup>OE</sup>) lentiviral vectors were constructed to express different transgene levels (Fig. 4b, c). ERDJ4<sup>OE</sup> conferred protection against tunicamycin-induced cell death in the TEX cell line as well as HSPCs (Fig. 4d and Extended Data Fig. 5c, d), suggesting that ERDJ4<sup>OE</sup> increases the threshold of ER stress needed to induce apoptosis. To test whether ERDJ4 influences human HSC function, transduced cord blood cells were transplanted into immune-deficient mice. ERDJ4<sup>OE</sup> transduced cells had a competitive advantage as compared to the control group (Fig. 4e and Extended Data Fig. 5e). To measure directly the impact of ERDJ4<sup>OE</sup> on the number of functional HSCs, *in vivo* limiting dilution analysis (LDA) was performed. At a low cell dose, ERDJ4<sup>OE</sup> resulted in higher engraftment compared to control transduction (Fig. 4f and Extended Data Fig. 5f, g). The LDA measurement demonstrated a 4.4-fold increase in the number of repopulating HSCs on ERDJ4<sup>OE</sup> (Fig. 4g). These data indicate that a protein folding factor classically associated with the UPR governs HSC function in xenograft assays.

To understand the mechanism of increased human HSC engraftment following ERDJ4<sup>OE</sup>, we found that ERDJ4<sup>OE</sup> does not change progenitor engraftment, the frequency of phenotypic stem and progenitor cells, lineage differentiation, homing or self-renewal as measured by secondary LDA (Fig. 4h and Extended Data Fig. 6a–d). We proposed that ERDJ4<sup>OE</sup> might protect against ER stress that could occur during *in vivo* transplantation. Gene expression analysis of control-transduced cord blood cells indicated increased CHOP and GADD34 expression after transplantation, consistent with a stress response (Fig. 4i). With ERDJ4<sup>OE</sup>, this surge in CHOP and GADD34 expression was absent, indicating that ERDJ4<sup>OE</sup> prevents upregulation of stress-related genes. Transplantation of human HSCs in the xenograft environment places them under replicative stress, which causes elevated ROS, DNA damage, and loss of HSC function<sup>2</sup>. These processes may be connected to the UPR as ROS and DNA damage can cause ER stress<sup>5</sup> and ROS accumulation leads to UPR-mediated apoptosis of HSCs<sup>24</sup>. The association between ER protein folding and HSC engraftment indicates that moderation of UPR activation may improve HSC survival during stem-cell transplantation.

Our results establish a previously unrecognized link between UPR signalling and human HSC function. Owing to distinct activation of UPR branches upon stress exposure, HSCs are rapidly cleared while progenitors are spared. This response of human HSCs is consistent with the selective induction of apoptosis after DNA damage or ROS accumulation<sup>2,25</sup>. Collectively, these experimental observations suggest that HSCs possess an intrinsic biological focus on preventing propagation following damage, reducing malignancy risk. Because terminal differentiation purges damaged progenitor cell progeny, clonal purity may be of less importance to progenitors. Loss of HSCs and intestinal stem cells in *Grp78*<sup>-/-</sup> mice suggests that stem cells of multiple tissues can interrogate ER stress and use differential UPR activation to mitigate against potentially pathological damage<sup>26,27</sup>. Overall, our data point to the elimination of individual



**Figure 4 | ERDJ4<sup>OE</sup> protects from tunicamycin-induced apoptosis and increases HSC output and frequency.** **a**, ERDJ4 expression in sorted cord blood populations (Extended Data Table 1). *P* values were calculated in comparison to CD49f<sup>+</sup> HSCs; qPCR results are shown as mean  $\pm$  s.e.m. of *n* = 3 cord blood samples. B/NK, B and NK cell progenitor. **b**, **c**, Validation of lentiviral vectors for ERDJ4<sup>OE</sup>. **b**, Transduced cord blood cells were analysed by qPCR. PGK and SFFV refer to lentiviral promoter driving ERDJ4 expression. Results are shown as mean  $\pm$  s.e.m. of *n* = 2 cord blood samples. **c**, Transduced K562 cells were analysed by western blot. ERK2 is shown as a loading control. **d**, ERDJ4<sup>OE</sup> protects from tunicamycin-induced apoptosis. Transduced HSPCs were treated with tunicamycin. Symbols represent *n* = 11 cord blood samples where control and ERDJ4<sup>OE</sup> groups are connected by a grey line; black line indicates mean  $\pm$  s.e.m.; *P* value was calculated using a paired *t*-test. **e**, ERDJ4<sup>OE</sup> confers a competitive advantage *in vivo*. Engraftment of transduced cord blood cells was analysed 20 weeks after injection. Data are

shown as mean  $\pm$  s.e.m. of *n* = 3 cord blood samples with 5 mice per group. **f**, **g**, ERDJ4<sup>OE</sup> increases the number of engrafting HSCs. **f**, Engraftment of transduced cord blood cells was analysed 10 weeks after injection of different cell doses. *P* value was calculated using the Mann–Whitney *U*-test. Every symbol represents one mouse; data represent *n* = 3 cord blood samples with 4 mice per group; line shows median. **g**, HSC frequency was calculated on the basis of the number of engrafted mice. **h**, Progenitors retain normal engraftment capacity with ERDJ4<sup>OE</sup>. Engraftment of transduced progenitors was assessed 2 and 4 weeks after injection. Data represents *n* = 3 cord blood samples with 3 mice per group; line shows median. **i**, ERDJ4<sup>OE</sup> moderates a surge in CHOP and GADD34 expression after transplantation. Transduced cord blood cells were expanded *in vitro* and transplanted into mice. CHOP and GADD34 expression in sorted GFP<sup>+</sup> cells was analysed by qPCR. Data are shown as mean  $\pm$  s.e.m. of *n* = 3 cord blood samples. \**P* < 0.05, \*\**P* < 0.01, \*\*\*\**P* < 0.0001.

HSCs after stress and damage as a paradigm of how the stem-cell pool maintains integrity, thereby ensuring long-term tissue maintenance.

## METHODS SUMMARY

Umbilical cord blood cells were enriched for CD34<sup>+</sup> cells and sorted by fluorescence-activated cell sorting (FACS), cultured with small molecules, and/or transduced with lentivirus. Quantitative RT–PCR was performed using primer sequences in Extended Data Table 3. Apoptosis was assessed by Annexin-V/Sytox staining followed by flow cytometry. Human HSC repopulation was read out by intrafemoral transplantation of cord blood cells into immune-deficient mice. Unless otherwise stated, *P* values were calculated by two-tailed unpaired Student's *t*-test.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 June 2013; accepted 5 March 2014.

Published online 28 April; corrected online 11 June 2014 (see full-text HTML version for details).

- Geiger, H., de Haan, G. & Florian, M. C. The ageing haematopoietic stem cell compartment. *Nature Rev. Immunol.* **13**, 376–389 (2013).
- Yahata, T. *et al.* Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. *Blood* **118**, 2941–2950 (2011).
- Shlush, L. I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328–333 (2014).

- Rossi, D. J., Jamieson, C. H. M. & Weissman, I. L. Stems cells and the pathways to aging and cancer. *Cell* **132**, 681–696 (2008).
- Rutkowski, D. T. & Kaufman, R. J. A trip to the ER: coping with stress. *Trends Cell Biol.* **14**, 20–28 (2004).
- Walter, P. & Ron, D. The unfolded protein response: from stress pathway to homeostatic regulation. *Science* **334**, 1081–1086 (2011).
- Majeti, R., Park, C. Y. & Weissman, I. L. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. *Cell Stem Cell* **1**, 635–645 (2007).
- Doulatov, S. *et al.* Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nature Immunol.* **11**, 585–593 (2010).
- Notta, F. *et al.* Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333**, 218–221 (2011).
- Laurenti, E. *et al.* The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. *Nature Immunol.* **14**, 756–763 (2013).
- DuRose, J. B., Tam, A. B. & Niwa, M. Intrinsic capacities of molecular sensors of the unfolded protein response to sense alternate forms of endoplasmic reticulum stress. *Mol. Biol. Cell* **17**, 3095–3107 (2006).
- Tabas, I. & Ron, D. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nature Cell Biol.* **13**, 184–190 (2011).
- Lu, P. D., Harding, H. P. & Ron, D. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J. Cell Biol.* **167**, 27–33 (2004).
- Vattem, K. M. & Wek, R. C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 11269–11274 (2004).
- Palam, L. R., Baird, T. D. & Wek, R. C. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J. Biol. Chem.* **286**, 10939–10949 (2011).

16. Marciniak, S. J. *et al.* CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum. *Genes Dev.* **18**, 3066–3077 (2004).
17. Han, J. *et al.* ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. *Nature Cell Biol.* **15**, 481–490 (2013).
18. Novoa, I., Zeng, H., Harding, H. P. & Ron, D. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2 $\alpha$ . *J. Cell Biol.* **153**, 1011–1022 (2001).
19. Boyce, M. *et al.* A selective inhibitor of eIF2 $\alpha$  dephosphorylation protects cells from ER stress. *Science* **307**, 935–939 (2005).
20. Axten, J. M. *et al.* Discovery of 7-methyl-5-(1-[[3-(trifluoromethyl)phenyl]acetyl]-2,3-dihydro-1H-indol-5-yl)-7H-pyrrolo[2,3-d]pyrimidin-4-amine (GSK2606414), a potent and selective first-in-class inhibitor of protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK). *J. Med. Chem.* **55**, 7193–7207 (2012).
21. Shen, Y., Meunier, L. & Hendershot, L. M. Identification and characterization of a novel endoplasmic reticulum (ER) DnaJ homologue, which stimulates ATPase activity of BiP *in vitro* and is induced by ER stress. *J. Biol. Chem.* **277**, 15947–15956 (2002).
22. Lai, C. W., Otero, J. H., Hendershot, L. M. & Snapp, E. ERdj4 protein is a soluble endoplasmic reticulum (ER) DnaJ family protein that interacts with ER-associated degradation machinery. *J. Biol. Chem.* **287**, 7969–7978 (2012).
23. Kurisu, J. *et al.* MDG1/ERdj4, an ER-resident DnaJ family member, suppresses cell death induced by ER stress. *Genes Cells* **8**, 189–202 (2003).
24. Rouault-Pierre, K. *et al.* HIF-2 $\alpha$  protects human hematopoietic stem/progenitors and acute myeloid leukemic cells from apoptosis induced by endoplasmic reticulum stress. *Cell Stem Cell* **13**, 549–563 (2013).
25. Milyavsky, M. *et al.* A distinctive DNA damage response in human hematopoietic Stem cells reveals an apoptosis-independent role for p53 in self-renewal. *Cell Stem Cell* **7**, 186–197 (2010).
26. Heijmans, J. *et al.* ER stress causes rapid loss of intestinal epithelial stemness through activation of the unfolded protein response. *Cell Rep.* **3**, 1128–1139 (2013).
27. Wey, S., Luo, B. & Lee, A. S. Acute inducible ablation of GRP78 reveals its role in hematopoietic stem cell survival, lymphogenesis and regulation of stress signaling. *PLoS ONE* **7**, e39047 (2012).

**Acknowledgements** We thank D. Ron and D. Rubinsztein for critical assessment of this work; all members of the Dick laboratory, especially O. I. Gan and E. Lechman, for experimental support and advice; A. Khandani, P. Penttilä and the SickKids-UHN flow facility for technical support; N. Jamal for providing bone marrow samples; J. Heijmans for providing the pLV-ca-GADD34 vector; and L. Naldini for providing the MA1 vector. Work in the Dick laboratory is supported by grants from the Canadian Institutes for Health Research, Canadian Cancer Society, Terry Fox Foundation, Genome Canada through the Ontario Genomics Institute, Ontario Institute for Cancer Research with funds from the province of Ontario, a Canada Research Chair and the Ontario Ministry of Health and Long Term Care (OMOHLTC). The views expressed do not necessarily reflect those of the OMOHLTC. A.R.G. and D.G.K. are supported by Leukemia and Lymphoma Research, Cancer Research UK, the Kay Kendall Leukaemia Fund, the NIHR Cambridge Biomedical Research Centre, the Cambridge Experimental Cancer Medicine Centre, and the Leukemia & Lymphoma Society of America. J.C.G. and T.F. are supported by Arthritis Research UK. S.J.M. and J.E.C. are supported by the Medical Research Council (UK). B.G.W. is supported by the Terry Fox New Frontiers Research Program (PPG09-020005), the Ontario Institute for Cancer Research, and the Canadian Institute for Health Research (CIHR grant 201592).

**Author Contributions** P.v.G., A.K. and J.E.D. designed the study and analysed and interpreted the data. P.v.G., A.K., N.M., D.G.K., T.F., J.E.C., S.X., K.H. and E.W. performed experiments. E.L. and K.E. performed bioinformatic analyses. S.J.M., J.C.G., A.R.G. and B.G.W. supervised specific experiments. P.v.G. wrote the paper. A.K. and J.E.D. revised the paper. J.E.D. supervised the study.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.E.D. ([jdick@uhnresearch.ca](mailto:jdick@uhnresearch.ca)).



## METHODS

**Cord blood and bone marrow sample preparation and liquid cell culture.** Umbilical cord blood, bone marrow or mobilized peripheral blood from healthy individuals was obtained with informed consent according to procedures approved by the institutional review boards of the University Health Network and Trillium Hospital. Mononuclear cells were obtained by centrifugation on Ficoll, and lineage depletion was performed using the StemSep Human Progenitor Cell Enrichment Kit according to the manufacturer's protocol (50–75% CD34<sup>+</sup>, StemCell Technologies). The Lin<sup>−</sup> cord blood or Lin<sup>−</sup> bone marrow/mobilized peripheral blood cells were stored in IMDM with 50% FCS and 10% DMSO at −150 °C until use. Unless stated otherwise, HSC/HSPC and progenitor fractions were sorted from cord blood. For liquid culture, cells were thawed and plated in X-VIVO 10 (BioWhittaker) supplemented with 1% BSA, 2 mM L-glutamine, 100 U ml<sup>−1</sup> penicillin-streptomycin, and the following cytokines: TPO (7.5 ng ml<sup>−1</sup>), SCF (50 ng ml<sup>−1</sup>), G-CSF (5 ng ml<sup>−1</sup>), Flt3 ligand (50 ng ml<sup>−1</sup>) and IL-6 (5 ng ml<sup>−1</sup>), referred to as TSGF6 culture. The cell line K562 was expanded in IMDM with 10% FCS, 2 mM L-glutamine and 100 U ml<sup>−1</sup> penicillin-streptomycin; TEX cells were cultured as reported previously<sup>28</sup>. Both cell lines tested negative for mycoplasma.

**Fluorescence-activated cell sorting and flow cytometry.** To separate cell populations, cord blood or bone marrow cells were re-suspended at 10<sup>7</sup> cells per ml, stained with antibodies against surface markers in PBS with 2% FCS, washed and sorted on the BD FACS Aria or MoFlo, consistently yielding >95% purity. Analytical flow cytometry was performed using BD LSRII or BD Canto cytometer. Data was analysed with FlowJo software (Tree Star, Inc.).

**Annexin-V/Sytox and cleaved caspase-3 flow cytometry.** For Annexin-V/Sytox apoptosis analysis, cells were stained for surface markers and washed in PBS with 2% FCS, re-suspended in binding buffer (diluted 10 × in H<sub>2</sub>O, BD catalogue number 556454) with Annexin-V-APC (50 × dilution, BD 550474) and Sytox Blue (500 × dilution, Life Technologies S34857) and stained for 20 min at room temperature. Then the sample was diluted 5 × with binding buffer and cells were analysed by flow cytometry within 60 min. For cleaved caspase-3 analysis, cells were permeabilized for 30 min at room temperature with BD Perm 2 buffer (diluted 10 × in H<sub>2</sub>O, catalogue number 347692), washed in 2 ml PBS with 2% FCS and stained with PE-conjugated cleaved caspase-3 antibody from BD (catalogue number 561011) for 30 min at room temperature. Cells were washed again and re-suspended in BD Cytofix buffer (diluted 4 × in PBS, BD catalogue number 554655) before flow cytometry analysis.

**Quantitative RT-PCR.** RNA was extracted from 500–100,000 cells using TRIzol (Life Technologies) and re-suspended in water for cDNA synthesis using the SuperScript III or SuperScript VILO systems according to manufacturer's instructions (Life Technologies). For each qPCR reaction we added 2 × Power SYBR Green mix (Life Technologies), 133 nM forward primer and 133 nM reverse primer and RNase-free water up to a total volume of 12.5 µl. cDNA was diluted 6–20 × with RNase-free water and 2.5 µl was added for each reaction. The qPCR was performed using a 7900 HT Real-Time PCR system with SDS v2.3 software (Applied Biosystems) using standard settings: 50 °C for 2 min; 95 °C for 10 min; then 95 °C for 15 s and 60 °C for 1 min repeated for 40 cycles; then dissociation stage. Each assay was run in duplicate for technical variation. Arbitrary mRNA concentrations were calculated using the relative standard curve method. Gene expression levels were normalized to *GAPDH* except in Fig. 1b (normalized to the average of *GAPDH*, *ACTB* and *PBGD*) and Fig. 4a (normalized to the average of *GAPDH* and *ACTB*). To determine *XBP1* splicing levels, primers were designed that amplify all or only spliced *XBP1* mRNA, and spliced *XBP1* expression was divided by total *XBP1* expression (indicated as spliced/total *XBP1*). qPCR primer sequences are listed in Extended Data Table 3.

**Methylcellulose colony-forming assays.** Methylcellulose (StemCell Technologies MethoCult H4034) was supplemented with IL-6 (10 ng ml<sup>−1</sup>), Flt3 ligand (10 ng ml<sup>−1</sup>) and DMSO or tunicamycin (0.6 µg ml<sup>−1</sup>). Using the BD FACS Aria, 500 CD34<sup>+</sup> CD38<sup>−</sup> CD45RA<sup>−</sup> CD90<sup>+</sup> HSCs or 350 CD34<sup>+</sup> CD38<sup>+</sup> progenitors were deposited in 2.5 ml methylcellulose, and duplicate dishes were plated with 1 ml each (200 HSCs or 140 progenitors per dish). After 13 days, colonies were counted and classified based on morphological appearance.

**Lentiviral vectors.** The bidirectional lentiviral MA1 vector<sup>29</sup> was modified by replacing ΔNGFR with a *loxP*-flanked Gateway cassette (Life Technologies) to generate the destination vector pMAL. To generate PGK-ERD4<sup>OE</sup>, pMAL was recombined with the *ERD4* entry vector from the Mammalian Gene Collection through the PlasmID Repository at Harvard<sup>30</sup> (Clone ID HsCD00076069). To generate constitutively active (ca) GADD34<sup>OE</sup>, the *GADD34* fragment was amplified by PCR from the pLV-ca-GADD34 construct<sup>26</sup> using the forward primer 5'-CACCATTGGCCCA GTGTGCTGG-3' and the reverse primer 5'-TCACTGGGAAGGAAGAAGG-3', cloned into an entry vector using the pENTR Directional TOPO cloning kit (Life Technologies K2400-20), and this entry vector was recombined with pMAL. To generate SFFV-ERD4<sup>OE</sup>, the PGK promoter of PGK-ERD4<sup>OE</sup> was replaced with the stronger SFFV promoter<sup>31</sup>. Control (Ctrl) vectors expressed a humanized *Renilla*

luciferase gene or a Stuffer sequence derived from pLKO.1\_1.9Kb\_stuffer instead of *ERD4* and *ca-GADD34*. The ATF4 reporter lentivirus was made as follows: first, the PGK promoter in pMAL was replaced with the stronger SFFV promoter and GFP was replaced with TagBFP (Evrogen<sup>32</sup>) to generate the destination vector pSMALB. Next, the reporter fragments were amplified from ATF4.5: 5' ATF4.GFP (Addgene 21852), ATF4.12: 5' ATF4.uORF1<sup>AUA</sup>.GFP (Addgene 21859) and ATF4.14: 5' ATF4.uORF1&2<sup>AUA</sup>.GFP (Addgene 21861) using the forward primer 5'-C ACCAGCTTTCTGCTTGCTGTC-3' and the reverse primer 5'-TTACTTGTA CAGCTCGTC CATGC-3'. These fragments were cloned into entry vectors using the pENTR Directional TOPO cloning kit. Finally, the entry vectors were recombined with pSMALB to generate the bidirectional lentiviral reporter vectors pSMALB-ATF4.5rep (referred to as ATF4 reporter), pSMALB-ATF4.12rep and pSMALB-ATF4.14rep, respectively. In the ATF4 reporter, mRNA expression highly correlates between TagBFP and ATF4-GFP due to the bidirectional promoter<sup>29</sup>. To account for differences in basal translation between experimental conditions (such as HSPCs versus progenitors, high versus low transduction, DMSO versus tunicamycin treatment), we calculated the transgene ratio between GFP and TagBFP as a measure of reporter activity (TGR = GFP mean fluorescence intensity/TagBFP mean fluorescence intensity, Fig. 3b). Reporter fluorescence was measured 30 h after addition of tunicamycin. For *in vitro* experiments, transduced cord blood cells were cultured for 3 days before treatment to allow for gene expression to come up.

**Lentivirus production and primary cell transduction.** Viral particles were pseudotyped with the vesicular stomatitis virus G (VSVG) protein using the pMD.G vector and third-generation pMDLg/pRRE and pRSVRev vectors were used for packaging in 293T cells using calcium-phosphate transfection (Clontech catalogue number 631312). Lentivirus was concentrated 100 × by ultracentrifugation, resuspended in X-VIVO 10 (BioWhittaker) supplemented with 1% BSA and stored at −80 °C until use. Cord blood cells were thawed and plated in liquid TSGF6 culture with double the concentration of all cytokines, and lentiviral suspension was added at a multiplicity of infection of 5–20 in a total volume of 100 µl (96 well plate) or 400 µl (24 well plate). After 16 h, TSGF6 culture medium was added to expand the cells. Transduction was measured using flow cytometry for GFP or TagBFP after 72–96 h.

**Mouse xenotransplantation and human Lin<sup>−</sup> cell isolation.** Mouse xenografts were performed as described previously<sup>33</sup> according to protocols approved by the University Health Network Animal Care Committee. Briefly, 8–12-week-old male NOD/Lt-scid/IL2Rγnull (NSG) mice were sublethally irradiated (225 cGy) 1 day before injection. Cells were injected intrafemorally with 30 µl PBS. Peripheral blood (80 µl) was taken from the saphenous vein and analysed by flow cytometry. Mice with >40% human T cells in the peripheral blood are likely to harbour an auto-immune clonal T-cell expansion and were excluded from analysis. To reach statistical significance, all animal studies were repeated with 3 cord blood samples with at least 2, but generally 5, mice per condition. Animals were not randomized before injection, no blinding was done for animal studies. After the mice were euthanized, femurs were flushed with 2 ml PBS 2% FCS; 50 µl was stained for surface markers and analysed by flow cytometry. Unless stated otherwise, engraftment data for injected femur is shown. For lineage depletion, bone marrow cells from the femurs and tibias of 5 mice were combined and processed with the StemSep Mouse/Human Chimera Enrichment kit (Stem Cell Technologies) according to the manufacturer's instructions. However, during the antibiotin incubation step, an additional 50 µl ml<sup>−1</sup> human haematopoietic progenitor enrichment antibody cocktail from the StemSep Human Progenitor Cell Enrichment Kit was added to deplete human lineage-positive cells.

**In vivo repopulation and serial limiting dilution analysis.** For competitive repopulation experiments, cord blood cells were transduced and ~50,000 cells were injected per mouse the next day. After 20 weeks, the mouse bone marrow was analysed for human CD45<sup>+</sup> engraftment as well as GFP and lineage differentiation. The median percentage of GFP<sup>+</sup> cells within the human CD45<sup>+</sup> graft was normalized to initial transduction for PGK-Ctrl and PGK-ERD4<sup>OE</sup> groups. For limiting dilution analysis (LDA) in primary recipients, transduced cord blood cells were expanded for 10 days in liquid TSGF6 culture and GFP<sup>+</sup> cells were sorted and injected at low and high cell doses. Human CD45<sup>+</sup> GFP<sup>+</sup> engraftment was analysed 10 weeks after transplantation. LDA in secondary recipient mice was performed by sorting GFP<sup>+</sup> cells from pooled bone marrow of primary mice. Secondary mice were injected with 30,000 to 1,000,000 cells and analysed 10 weeks after transplantation; a secondary mouse was scored as positive if it had >0.01% human engraftment. HSC frequency was estimated using the ELDA software<sup>34</sup> (Extreme Limiting Dilution Analysis, <http://bioinf.wehi.edu.au/software/elda/>). Progenitor engraftment was tested by overnight transduction of sorted CD34<sup>+</sup> CD38<sup>−</sup> progenitors and injection of 8,700–14,000 cells per mouse. The percentage of human CD45<sup>+</sup> GFP<sup>+</sup> cells in the mouse bone marrow after 2–4 weeks was normalized to initial transduction. To assess gene expression changes on transplantation, mice were killed at 19 h, 1 week and 2 weeks after transplantation of 1.0–1.6 × 10<sup>6</sup> transduced cord blood

cells. Human GFP<sup>+</sup>CD34<sup>+</sup> cells were sorted from the xenografted mouse bone marrow (at 19 h and 1 week after transplant, all GFP<sup>+</sup> cells were sorted due to cell number constraints).

**Analysis of *Chop*<sup>-/-</sup> mouse bone marrow.** *Chop*<sup>-/-</sup> and wild-type littermates on the C57BL/6 background were housed in a specific pathogen-free facility according to the Animals Scientific Procedures outlined by the UK. Femurs, tibias and pelvises of 13-week-old male mice were flushed to isolate bone marrow cells. Five to ten million cells were stained with antibodies against surface markers (HSC panel: lineage markers, c-Kit, Sca-1, CD34 and Flk2; progenitor panel: lineage markers, c-Kit, Sca-1, CD34 and FcγII/IIIR, Extended Data Table 2). Next, cells were stained using Annexin-V Pacific Blue conjugate (Biolegend) and 7-AAD according to manufacturer's instructions, and analysed on a BD Fortessa cytometer.

**Tunicamycin, thapsigargin, salubrinal and GSK2606414.** Compounds were purchased as follows: tunicamycin, Sigma-Aldrich, catalogue number T7765; thapsigargin, Sigma-Aldrich, catalogue number T9033; salubrinal, Santa Cruz, catalogue number SC202332; GSK2606414, Merck Millipore, catalogue number 516535. Powder was re-suspended in DMSO and stored at -20 °C until use. Final DMSO concentration was always <0.6% and equal between treatment and control groups. Unless otherwise indicated, cell counts and viability analyses were performed after 40 h of treatment. Viable cells were counted manually by Trypan blue exclusion or automated using the BD Canto flow cytometer high throughput sampler (HTS), by counting the number of Annexin<sup>-</sup> and/or Sytox<sup>-</sup> cells in a specified volume.

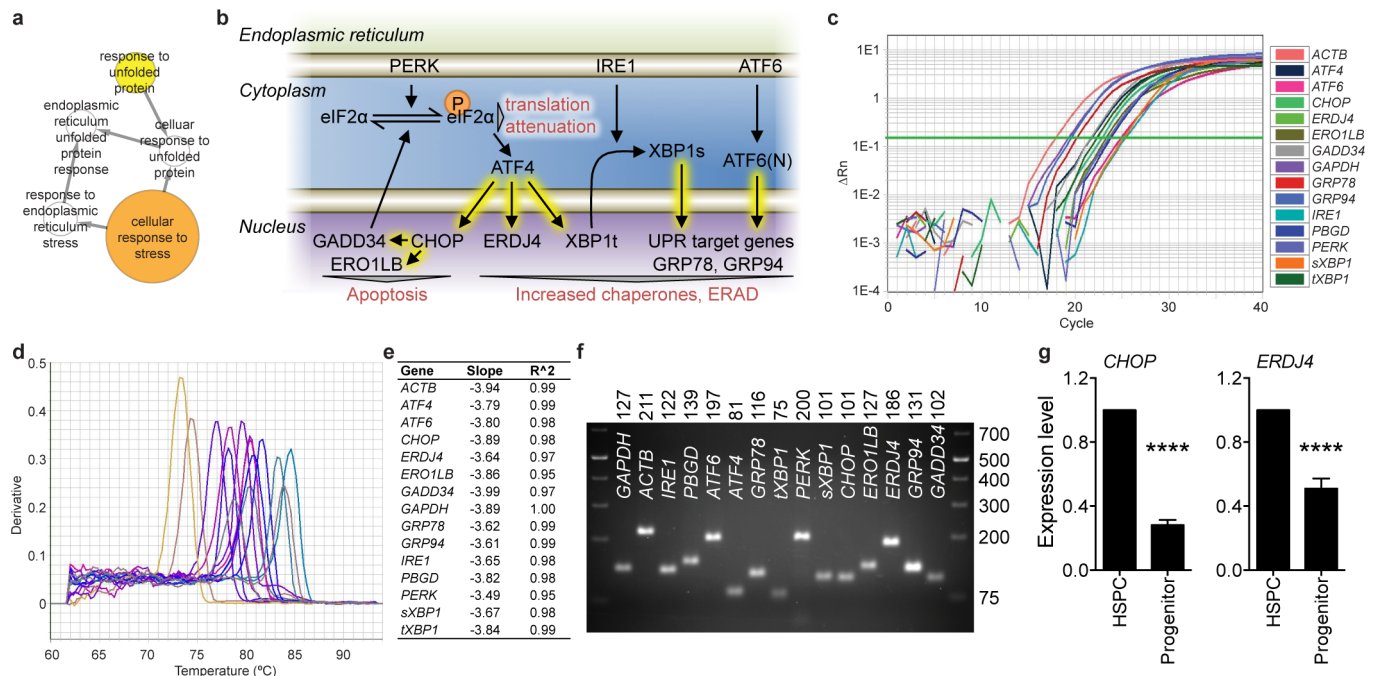
**Western blot.** Transduced K562 cells were lysed, separated with SDS-PAGE and transferred onto a polyvinylidene fluoride membrane as previously reported<sup>33</sup>. Specific antibody to ERDJ4 (Abnova catalogue number H00004189-M09) was detected using secondary HRP-conjugated antibodies (Amersham) and visualized by chemiluminescence (Pierce).

**Gene expression and pathway analysis.** Gene expression data sets were reported previously<sup>10</sup>. The genes upregulated in HSC compared to progenitor (MLP/CMP/GMP/MEP) with adjusted *P* value <0.01 were used for gene ontology (GO) analyses with BiNGO<sup>35</sup>. The algorithm was used with hypergeometric test, multiple test correction (Benjamini-Hochberg false discovery rate (FDR)) and using the whole *Homo sapiens* annotation as a reference set. Data were visualized with Cytoscape<sup>36</sup>; gene sets linked to the UPR response were unlinked from the rest of the network for presentation. For the heat map, we checked the expression of all genes belonging to GO categories relative to the UPR response (79 genes belonging to GO\_0006986,

GO\_0034620, GO\_0034976, GO\_0030968 categories). Forty of these were differentially expressed between HSC and progenitors (FDR <0.05). Their expression levels in HSC, MPP, MLP, CMP, GMP and MEP were mean centred.

**Statistical analysis.** Unless otherwise stated, mean ± s.e.m. values are given and *P* values were calculated by two-tailed unpaired Student's *t*-test. Mann-Whitney *U*-tests were performed to compare engraftment levels, as these data do not show a normal distribution. ELDA software was used for statistical analysis of *in vivo* LDA<sup>34</sup> (<http://bioinf.wehi.edu.au/software/elda>). \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001.

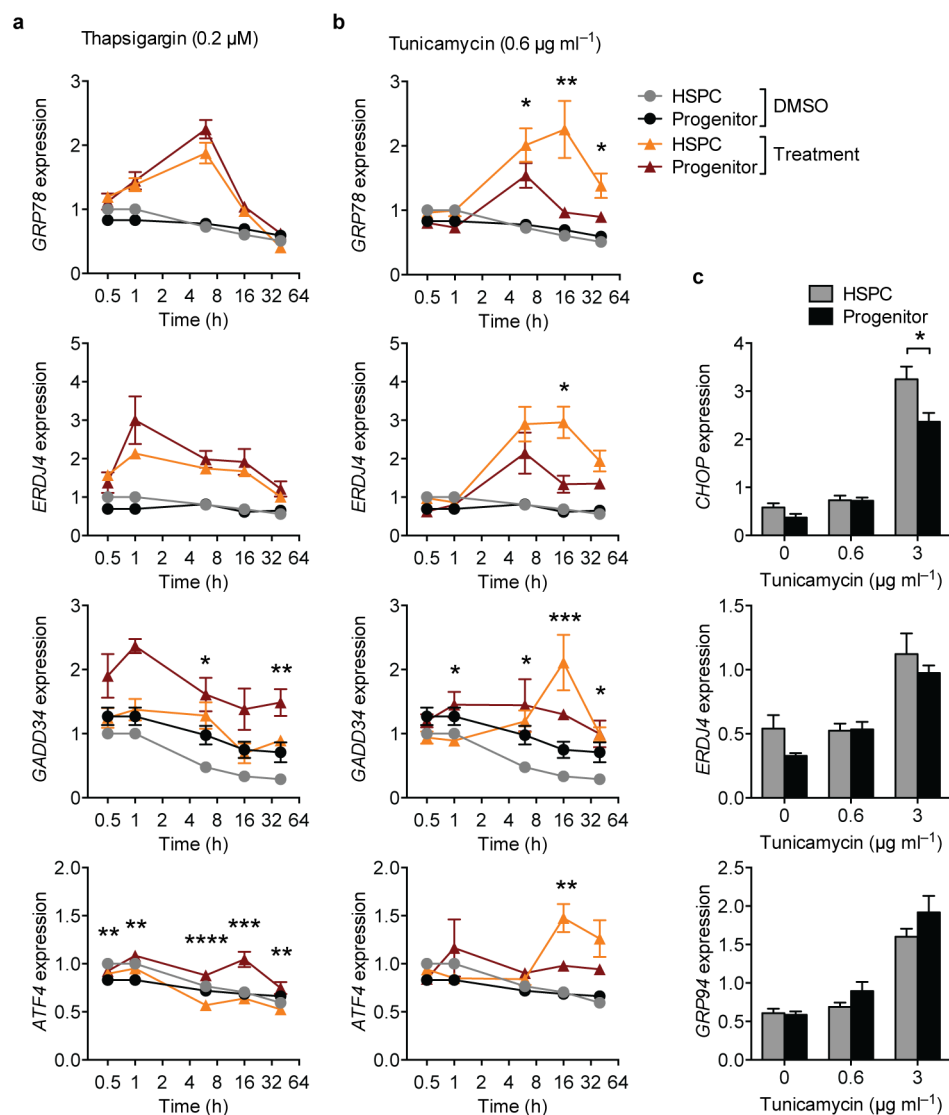
28. Warner, J. K. *et al.* Direct evidence for cooperating genetic events in the leukemic transformation of normal human hematopoietic cells. *Leukemia* **19**, 1794–1805 (2005).
29. Amendola, M., Venneri, M. A., Biffi, A., Vigna, E. & Naldini, L. Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters. *Nature Biotechnol.* **23**, 108–116 (2005).
30. The MGC Project Team. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res.* **14**, 2121–2127 (2004).
31. Gentner, B. *et al.* Stable knockdown of microRNA *in vivo* by lentiviral vectors. *Nature Methods* **6**, 63–66 (2009).
32. Subach, O. M. *et al.* Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **15**, 1116–1124 (2008).
33. van Galen, P. *et al.* Reduced lymphoid lineage priming promotes human hematopoietic stem cell expansion. *Cell Stem Cell* **14**, 94–106 (2014).
34. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 (2009).
35. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).
36. Lopes, C. T. *et al.* Cytoscape Web: an interactive web-based network browser. *Bioinformatics* **26**, 2347–2348 (2010).
37. Yoshida, H., Matsui, T., Yamamoto, A., Okada, T. & Mori, K. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* **107**, 881–891 (2001).
38. Lin, J. H. *et al.* IRE1 signaling affects cell fate during the unfolded protein response. *Science* **318**, 944–949 (2007).
39. Harding, H. P., Zhang, Y. & Ron, D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* **397**, 271–274 (1999).



### Extended Data Figure 1 | Expression analysis of UPR-related genes.

**a**, Enrichment of UPR-related genes in human HSCs compared to progenitors. CD49f<sup>+</sup> HSC-enriched genes were analysed for GO category overrepresentation. Node size represents the number of genes; white, yellow and orange colour correspond to FDR < 0.15, < 0.1 and < 0.01. **b**, Simplified scheme illustrating UPR signalling events. Three branches of the UPR are activated upon ER stress: IRE1, PERK and ATF6. IRE1 splices cytosolic *XBP1* mRNA to enable translation of the XBP1s transcription factor, which upregulates chaperones and ER-associated degradation (ERAD) machinery to resolve ER stress<sup>37,38</sup>. PERK initiates a different branch of the UPR through phosphorylation of eIF2 $\alpha$ , which attenuates global protein synthesis, thus permitting time to restore ER homeostasis<sup>39</sup>. Prolonged ER stress leads to PERK signalling-mediated upregulation of the proapoptotic transcription factor *CHOP* and its target *GADD34*. *GADD34* dephosphorylates eIF2 $\alpha$  leading to restoration of global protein translation. However, if ER stress is not resolved, *GADD34* upregulation can lead to further accumulation of misfolded proteins, oxidative stress and apoptosis<sup>16</sup>. Yellow highlighted arrows indicate

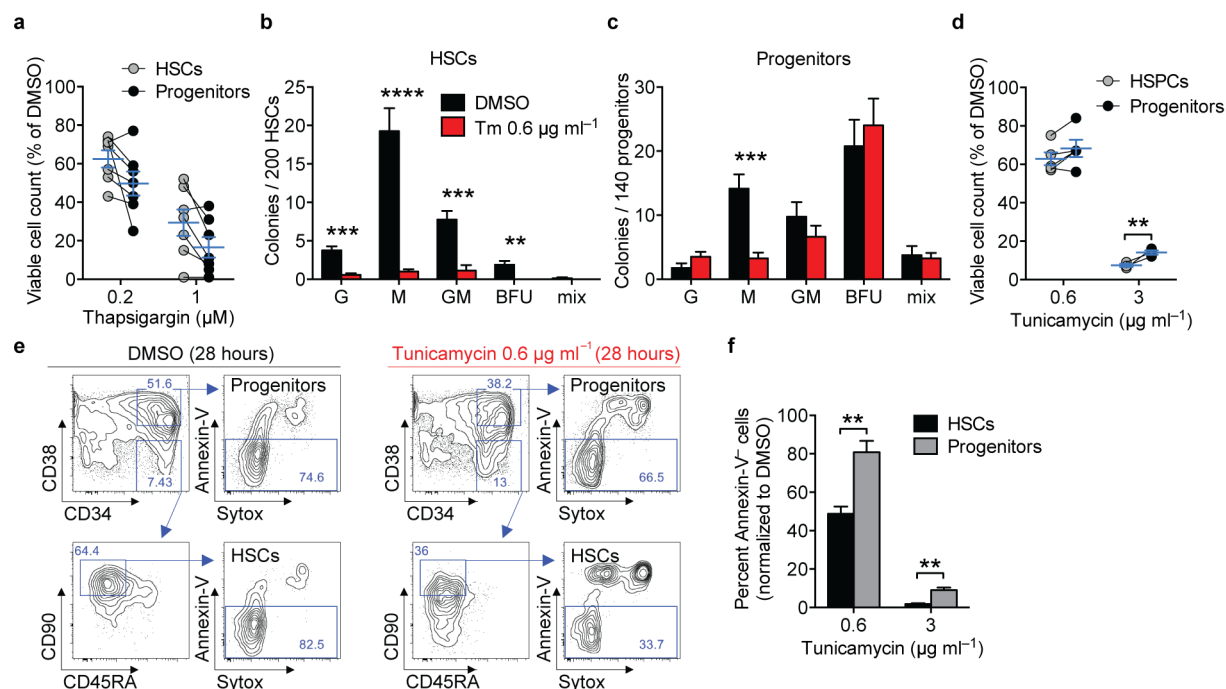
transcriptional regulation. **c**, Amplification curves of qPCR reactions for UPR-related genes. Fluorescence signal during 40 cycles of qPCR reactions on cord-blood-derived cDNA is shown for a representative experiment. Green line indicates threshold that was used to calculate mRNA quantity. **d**, Dissociation curves were generated to check for the presence of aspecific amplicons or primer dimers, which would be visible as additional peaks. Each line represents the dissociation curve of one qPCR reaction, colours indicate different genes. **e**, Slopes and  $R^2$  values of standard curves are shown for a representative experiment. These values were calculated separately for each experiment, based on a cDNA dilution series. **c–e** were performed using SDS v2.3 software. **f**, Agarose gel analysis of qPCR amplicons. qPCR reactions were run on a 3% agarose gel to check for reaction specificity: nonspecific amplicons would be visible as additional bands. The expected product size is shown above the gel; the ladder sizes are indicated on the right. **g**, Adult bone marrow cells were sorted into HSPC and progenitor fractions. mRNA levels for *CHOP* and *ERDJ4* were measured by qPCR. Results are shown as mean  $\pm$  s.e.m. of  $n = 5$  bone marrow samples. \*\*\*\* $P < 0.0001$ .



**Extended Data Figure 2 | Differential response of HSPCs and progenitors to ER-stress-inducing agents.** **a, b**, HSPC and progenitor fractions were sorted and plated in the presence of **(a)** thapsigargin or **(b)** tunicamycin. mRNA was isolated after 0.5, 1, 6, 16 and 40 h and expression levels of *GRP78*, *ERDJ4*, *GADD34* and *ATF4* were assessed by qPCR. The DMSO-treated controls were the same between **a** and **b**. Data are shown as mean  $\pm$  s.e.m. of  $n = 3$  cord

blood samples,  $P$  value was calculated based on fold change of treated over DMSO control cells and indicates differential response between HSPCs and progenitors. **c**, Adult bone marrow HSPCs and progenitors were sorted and plated in the presence of tunicamycin. After 16 h, mRNA was isolated and expression levels of *CHOP*, *ERDJ4* and *GRP94* were assessed by qPCR. Data are shown as mean  $\pm$  s.e.m. of  $n = 5$  bone marrow samples.

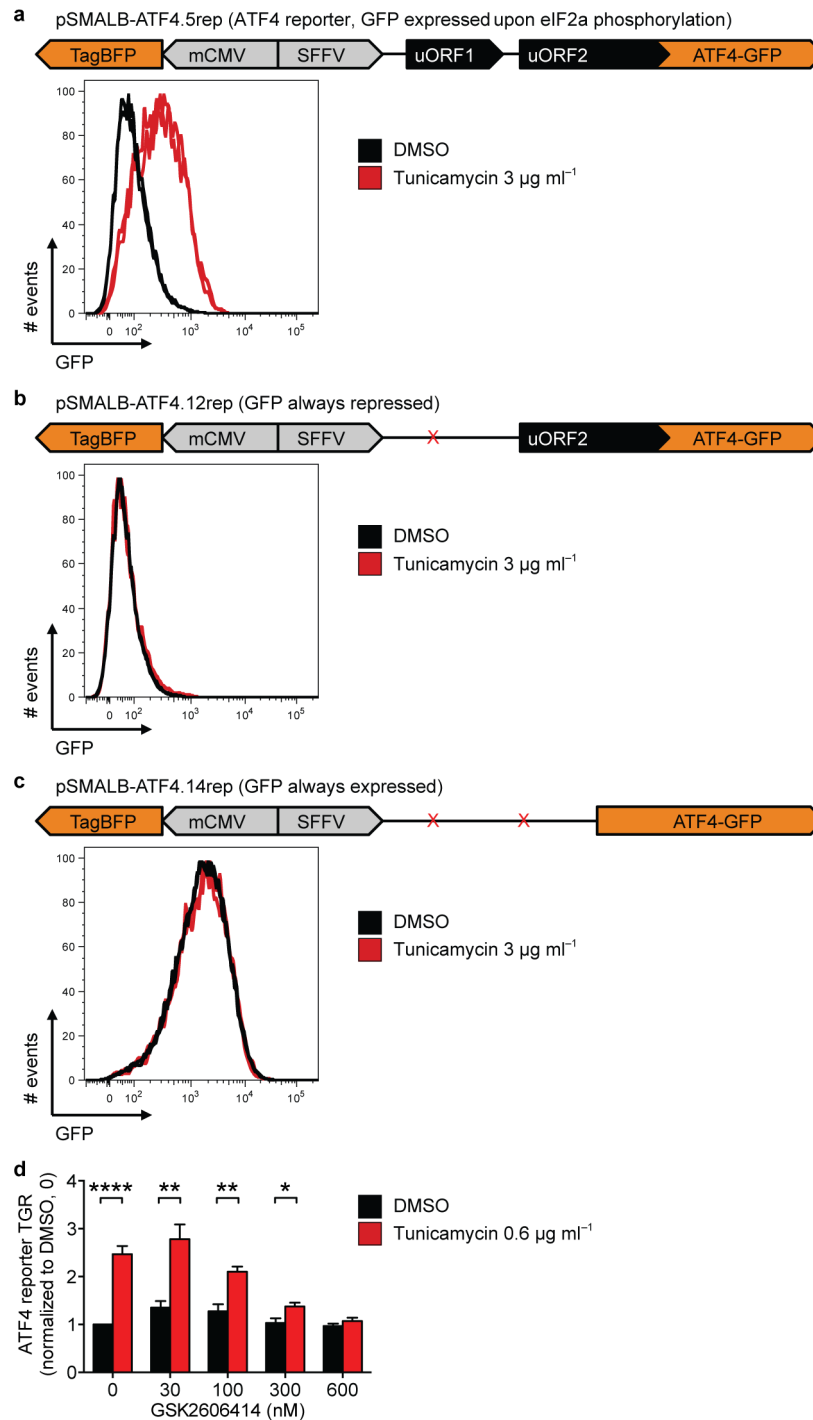




### Extended Data Figure 3 | Survival of HSCs is lower compared to progenitors after tunicamycin, but not thapsigargin treatment.

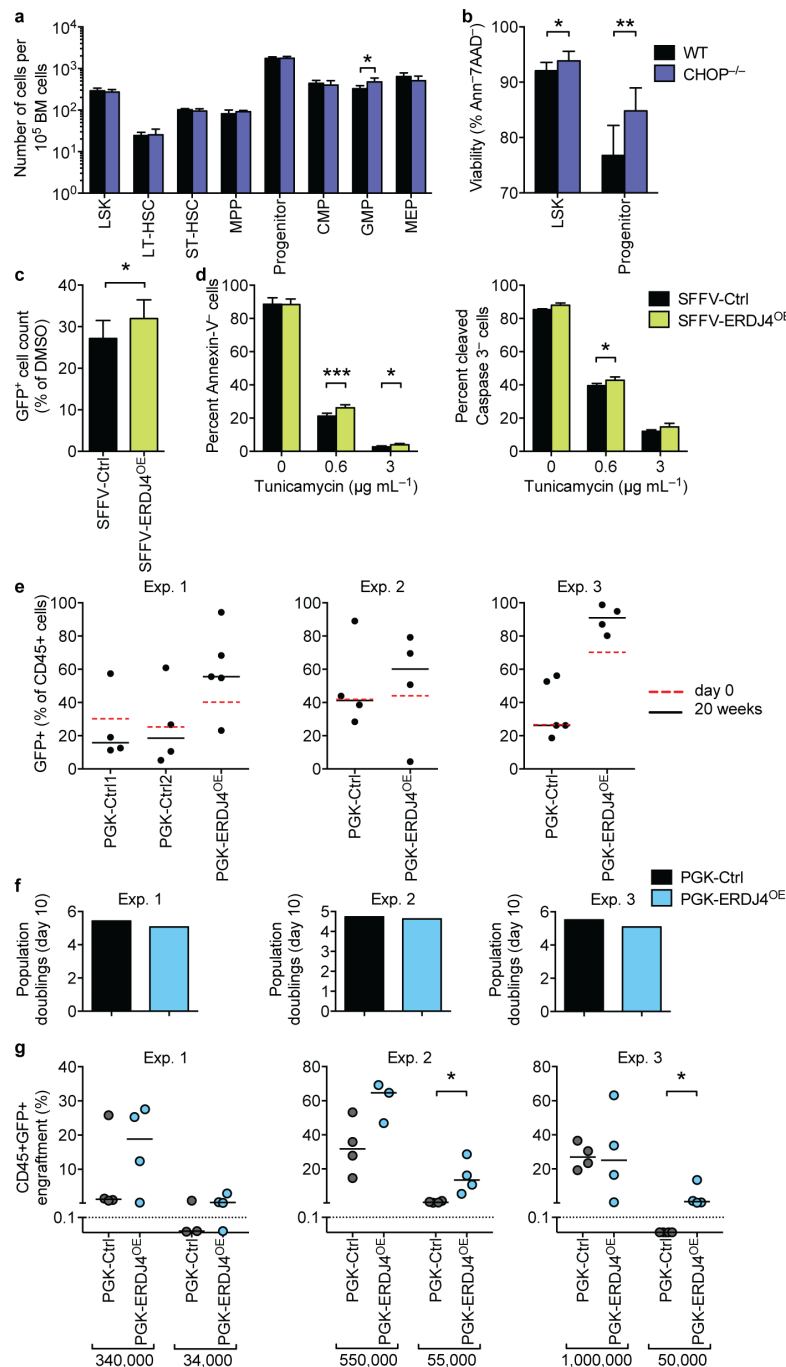
**a**, Thapsigargin has similar toxicity for sorted HSC and progenitor fractions. Sorted HSCs and progenitors were plated in TSGF6 culture conditions in the presence of thapsigargin or DMSO control. Symbols represent viable cell counts of individual samples where fractions are connected by a black line; the blue line indicates mean  $\pm$  s.e.m. of  $n = 7$  cord blood samples. **b**, **c**, Reduced clonogenic capacity of sorted HSCs compared to progenitors after tunicamycin treatment. Total colony number is shown in Fig. 2c. Here, data are separated into colony types based on morphological appearance. Data are shown as mean  $\pm$  s.e.m. of  $n = 4$  cord blood samples. G, granulocyte; M, macrophage; GM, granulocyte/macrophage; BFU, erythroid burst forming unit; mix, multilineage. **d**, HSCs have lower survival compared to progenitors after tunicamycin treatment, even after cell cycle induction. Sorted HSC and

progenitor fractions were plated in TSGF6 culture conditions with double cytokine concentrations for 72–96 h to induce G0 exit of the HSC fraction<sup>33</sup>. Then, cells were plated in the presence of tunicamycin. Viable cell counts as a percentage of DMSO controls are shown. Symbols represent individual samples where fractions are connected by a black line; the blue line indicates mean  $\pm$  s.e.m. of  $n = 5$  cord blood samples at  $0.6 \mu\text{g ml}^{-1}$  and  $n = 3$  cord blood samples at  $3 \mu\text{g ml}^{-1}$  tunicamycin. **e**, **f**, Increased apoptosis of HSCs compared to progenitors after tunicamycin treatment. **e**, Cord blood cells were plated with tunicamycin and stained for primitive surface markers, Annexin-V and Sytox. Representative flow plots are shown. **f**, Sorted HSCs and progenitors were plated in the presence of tunicamycin. The percentage of viable Annexin-V<sup>-</sup> cells after 40 h compared to DMSO controls is shown as mean  $\pm$  s.e.m. of  $n = 4$  cord blood samples. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .



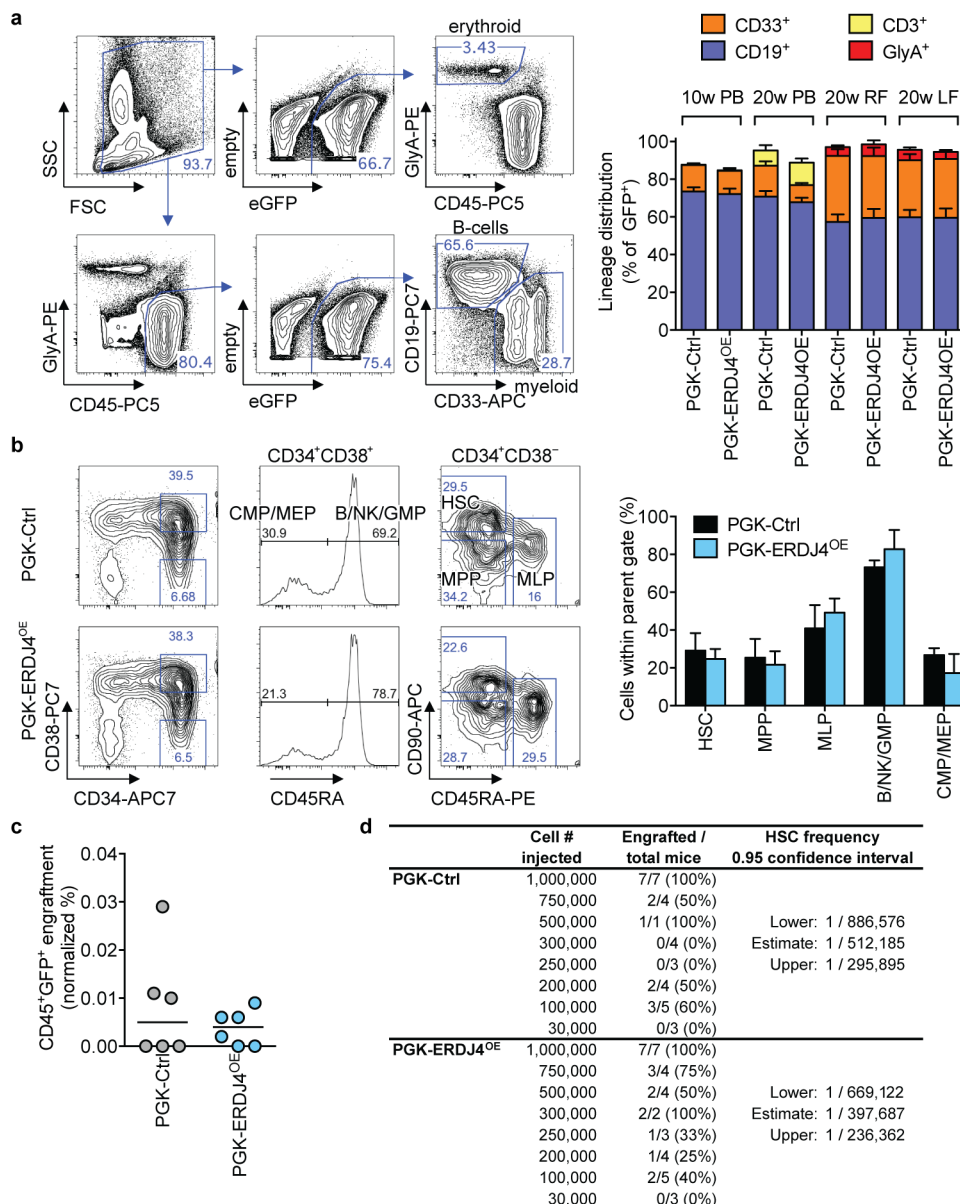
**Extended Data Figure 4 | ATF4 reporter enables visualization of increased ATF4 translation after tunicamycin treatment.** **a**, ATF4 reporter validation. Two upstream ORFs (uORFs) that are 5' of the *ATF4* coding sequence in the *ATF4* mRNA ensure more efficient translation of ATF4 when eIF2 $\alpha$  phosphorylation levels are high<sup>13,14</sup>. A bidirectional lentiviral vector was constructed that gives constitutive expression of TagBFP to mark transduced cells. In the other direction, the SFFV promoter drives expression of the 5' end of the *ATF4* mRNA which fuses with a GFP reporter gene 3' of the termination codon of uORF2. HeLa cells were transduced with pSMALB-ATF4.5rep (referred to as ATF4 reporter) and treated with tunicamycin. After 30 h, GFP fluorescence was read out by flow cytometry. Histogram plots show  $n = 2$

technical duplicates (two black lines for DMSO control, two red lines for tunicamycin treatment). **b**, **c**, Reporter fluorescence depends on uORFs. HeLa cells were transduced and treated with tunicamycin. As expected, ATF4-GFP translation is (**b**) repressed in the negative control that has a mutated uORF1 start codon and (**c**) constitutively high in the positive control with mutated start codons for both uORFs<sup>13</sup>. Histogram plots show  $n = 2$  technical duplicates. **d**, ATF4 reporter-transduced cord blood cells were treated with tunicamycin and increasing doses of the PERK inhibitor GSK2606414. The transgene ratio (TGR) is shown as mean  $\pm$  s.e.m. of  $n = 6$  cord blood samples (except at 600 nM,  $n = 3$  cord blood samples). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ .



**Extended Data Figure 5 | Modulation of UPR-associated genes affects haematopoietic stem and progenitor cells *in vivo*.** **a**, Analysis of haematopoietic stem and progenitor cell frequencies in *Chop*<sup>-/-</sup> mice. Flow cytometry was performed on mouse bone marrow (Extended Data Table 2). Bars show the absolute cell production in each population from wild-type or *Chop*<sup>-/-</sup> mice. Data are shown as mean  $\pm$  s.d. of  $n = 5$  mice per group. **b**, Viability analysis of stem and progenitor cell populations in *Chop*<sup>-/-</sup> mice. The percentage of viable Annexin-V<sup>+</sup> 7-AAD<sup>-</sup> cells within the HSC-enriched LSK and Lin<sup>-</sup>Sca-1<sup>-</sup>c-Kit<sup>+</sup> progenitor fractions was assessed by flow cytometry. Data are shown as mean  $\pm$  s.d. of technical duplicates of  $n = 5$  mice per group. **c**, ERDJ4<sup>OE</sup> cells show increased survival after tunicamycin treatment. The haematopoietic TEX cell line<sup>28</sup> was transduced with SFFV-Ctrl or SFFV-ERDJ4<sup>OE</sup> lentiviral vectors and plated in the presence of 0.6  $\mu\text{g mL}^{-1}$  tunicamycin (SFFV refers to lentiviral promoter driving transgene expression). After 48 h, the number of transduced cells compared to DMSO-treated controls was determined by automated counting of GFP<sup>+</sup> cells. Data are shown as mean  $\pm$  s.d. of  $n = 3$  independent experiments,  $P$  value was calculated using a paired  $t$ -test. **d**, Tunicamycin-induced apoptosis is reduced by ERDJ4<sup>OE</sup>. Cells

from **c** were analysed for Annexin-V and cleaved caspase-3 expression by flow cytometry. Data are shown as mean  $\pm$  s.d. of  $n = 3$  independent experiments,  $P$  values were calculated using paired  $t$ -tests. **e**, ERDJ4<sup>OE</sup> endows cord blood cells with a competitive advantage over untransduced cells. Three cord blood pools (Exp. 1–3) were transduced with PGK-Ctrl or PGK-ERDJ4<sup>OE</sup> lentiviruses and injected into 5 mice each. Dashed line indicates GFP% after transduction (day 0); solid line indicates median GFP% of the human CD45<sup>+</sup> graft in the injected femur of xenografted mice (20 weeks). Every symbol represents one mouse. **f**, Similar expansion of PGK-Ctrl and PGK-ERDJ4<sup>OE</sup> transduced cord blood cells *in vitro*. Three cord blood pools (Exp. 1–3) were transduced with PGK-Ctrl or PGK-ERDJ4<sup>OE</sup> lentivirus and expanded for 10 days in liquid culture. Total population doublings of transduced GFP<sup>+</sup> cells is shown. **g**, ERDJ4<sup>OE</sup> increases HSC output. After liquid culture, GFP<sup>+</sup> cells from **f** were sorted and injected at high and low cell doses, indicated below the  $x$  axis. Total human CD45<sup>+</sup>GFP<sup>+</sup> engraftment in the injected femur after 10 weeks is shown.  $P$  values were calculated using the Mann–Whitney  $U$ -test. Every symbol represents one mouse, line shows median. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .



# Extended Data Figure 6 | Lineage differentiation, progenitor cell frequencies, homing, and serial transplantability are maintained following ERDJ4<sup>OE</sup>.

**a**, PGK-ERDJ4<sup>OE</sup>-transduced cord blood maintains multilineage differentiation potential *in vivo*. Left: gating scheme to assess differentiation of the human graft in mouse bone marrow. Representative flow plots show quantification of CD45<sup>+</sup>CD19<sup>+</sup> B cells, CD45<sup>+</sup>CD33<sup>+</sup> monocytes and granulocytes, and CD45<sup>+</sup>GlyA<sup>+</sup> erythroid cells within the GFP<sup>+</sup> graft. Right: the differentiation of transduced cord blood cells was assessed in the peripheral blood (PB) at 10 and 20 weeks and in the injected (RF) and non-injected (LF) femur at 20 weeks after transplantation. Results are shown as mean  $\pm$  s.e.m. of  $n = 15$  mice representing  $n = 3$  cord blood samples.

**b**, ERDJ4<sup>OE</sup> does not cause aberrant expansion of stem or progenitor cell fractions. To assess the distribution of human stem and progenitor cells, lineage<sup>+</sup> and mouse cells were depleted from xenografted mouse bone marrow. The remaining human lineage<sup>+</sup> cells were analysed by flow cytometry. Left: gating scheme to assess differentiation into HSC, MPP, MLP, CMP/MEP and B/NK/GMP fractions (Extended Data Table 1). Right: the frequency of human

stem and progenitor cells within the human CD45<sup>+</sup>GFP<sup>+</sup> graft was assessed 20 weeks after transplantation of transduced cord blood cells. Results are shown as mean  $\pm$  s.e.m. of  $n = 3$  cord blood samples. **c**, Homing capacity to the non-injected bone marrow is not altered by ERDJ4<sup>OE</sup>. Transduced cord blood cells were expanded for 12 days in liquid culture conditions and  $1-1.6 \times 10^6$  cells were transplanted per mouse. After 19 h, mice were euthanized to assess human CD45<sup>+</sup>GFP<sup>+</sup> cell homing to the non-injected femur. Results were normalized to transduction efficiency. Every symbol represents one mouse, results of  $n = 3$  cord blood samples are shown with 2 mice per group each; line shows median. **d**, Frequency of functional human HSCs *in vivo* is maintained with ERDJ4<sup>OE</sup>. Cord blood cells were transduced and injected into primary mice. After 10 weeks, mice were killed and transduced GFP<sup>+</sup> cells were sorted from their bone marrow. Thirty thousand to one million cells were re-transplanted into secondary mice for serial LDA. After 10 weeks, the bone marrow of secondary mice was assessed for human CD45<sup>+</sup>GFP<sup>+</sup> engraftment; mice were scored as positive if the engraftment level was  $>0.01\%$ . Data from  $n = 3$  cord blood samples was pooled.



Extended Data Table 1 | Surface marker phenotypes to separate human stem and progenitor cell subsets

Population	Surface phenotype
HSPC	CD34 <sup>+</sup> CD38 <sup>-</sup>
HSC	CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>-</sup> CD90 <sup>+</sup>
CD49f <sup>+</sup> HSC	CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>-</sup> CD90 <sup>+</sup> CD49f <sup>+</sup>
MPP	CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>-</sup> CD90 <sup>-</sup>
CD49f <sup>-</sup> MPP	CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>-</sup> CD90 <sup>-</sup> CD49f <sup>-</sup>
MLP	CD34 <sup>+</sup> CD38 <sup>-</sup> CD45RA <sup>+</sup> CD90 <sup>-/lo</sup>
Progenitor	CD34 <sup>+</sup> CD38 <sup>+</sup>
CMP/MEP	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>-</sup>
B/NK/GMP	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>+</sup>
CMP	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>-</sup> CD10 <sup>-</sup> CD135 <sup>+</sup>
MEP	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>-</sup> CD10 <sup>-</sup> CD135 <sup>-</sup>
B/NK	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>+</sup> CD10 <sup>+</sup>
GMP	CD34 <sup>+</sup> CD38 <sup>+</sup> CD45RA <sup>+</sup> CD10 <sup>-</sup> CD135 <sup>+</sup>

HSPC, haematopoietic stem and progenitor cell; HSC, haematopoietic stem cell; MPP, multipotent progenitor; MLP, multilymphoid progenitor; CMP, common myeloid progenitor; MEP, megakaryocyte erythrocyte progenitor; B/NK, B and NK cell progenitor; GMP, granulocyte macrophage progenitor. ‘HSPC’ to ‘HSC’ to ‘CD49f<sup>+</sup> HSC’ indicates increasing purity of the population: approximately 1:75, 1:20 and 1:10, respectively<sup>7,9</sup>.

Extended Data Table 2 | Surface marker phenotypes to separate mouse stem and progenitor cell subsets

Population	Surface phenotype
LSK	Lin <sup>−</sup> Sca-1 <sup>+</sup> c-Kit <sup>+</sup>
LT-HSC	Lin <sup>−</sup> Sca-1 <sup>+</sup> c-Kit <sup>+</sup> Flk2 <sup>−</sup> CD34 <sup>−</sup>
ST-HSC	Lin <sup>−</sup> Sca-1 <sup>+</sup> c-Kit <sup>+</sup> Flk2 <sup>−</sup> CD34 <sup>+</sup>
MPP	Lin <sup>−</sup> Sca-1 <sup>+</sup> c-Kit <sup>+</sup> Flk2 <sup>+</sup> CD34 <sup>+</sup>
Progenitor	Lin <sup>−</sup> Sca-1 <sup>−</sup> c-Kit <sup>+</sup>
CMP	Lin <sup>−</sup> Sca-1 <sup>−</sup> c-Kit <sup>+</sup> FcγII/IIIIR <sup>lo</sup> CD34 <sup>+</sup>
GMP	Lin <sup>−</sup> Sca-1 <sup>−</sup> c-Kit <sup>+</sup> FcγII/IIIIR <sup>hi</sup> CD34 <sup>+</sup>
MEP	Lin <sup>−</sup> Sca-1 <sup>−</sup> c-Kit <sup>+</sup> FcγII/IIIIR <sup>−</sup> CD34 <sup>−</sup>

LSK, HSC-enriched Lin<sup>−</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup> cells; LT-HSC, long-term haematopoietic stem cell; ST-HSC, short-term haematopoietic stem cell; MPP, multipotent progenitor; CMP, common myeloid progenitor; GMP, granulocyte macrophage progenitor; MEP, megakaryocyte erythrocyte progenitor.

**Extended Data Table 3 | Primer sequences used for quantitative RT-PCR**

Gene	Primer	Sequence 5' to 3'
<i>ACTB</i>	forward	CCTGTACGCCAACACAGTGC
<i>ACTB</i>	reverse	ATACTCCTGCTTGCTGATCC
<i>ATF4</i>	forward	GCTAAGGCGGGCTCCTCCGA
<i>ATF4</i>	reverse	ACCCAACAGGGCATCCAAGTCG
<i>ATF6</i>	forward	ATGAAGTTGTGTCAGAGAAACC
<i>ATF6</i>	reverse	CTCTTTAGCAGAAAATCCTAG
<i>CHOP/DDIT3</i>	forward	GGAGCATCAGTCCCCCACTT
<i>CHOP/DDIT3</i>	reverse	TGTGGGATTGAGGGTCACATC
<i>ERDJ4/DNAJB9</i>	forward	TCGGCATCAGAGCGCCAAATCA
<i>ERDJ4/DNAJB9</i>	reverse	ACCACTAGTAAAAGCACTGTGTCCAAG
<i>ERO1LB</i>	forward	TTCTGGATGATTGCTTGTGTGAT
<i>ERO1LB</i>	reverse	GGTCGCTTCAGATTAACCTTGT
<i>GADD34/PPP1R15A</i>	forward	CCCAGAAACCCCTACTCATGATC
<i>GADD34/PPP1R15A</i>	reverse	GCCCAGACAGCCAGGAAAT
<i>GRP78/HSPA5/BiP</i>	forward	TGACATTGAAGACTTCAAAGCT
<i>GRP78/HSPA5/BiP</i>	reverse	CTGCTGTATCCTCTTCAACAGT
<i>GRP94/HSP90B1/TRA1</i>	Qiagen	Qiagen cat. number QT00046963
<i>IRE1/ERN1</i>	forward	TGCTTAAGGACATGGCTACCATCA
<i>IRE1/ERN1</i>	reverse	CTGGAACTGCTGGTGCTGGA
<i>PERK/EIF2AK3</i>	forward	AATGCCTGGGACGTGGTGGC
<i>PERK/EIF2AK3</i>	reverse	TGGTGGTGCTTCGAGCCAGG
<i>PBGD/HMBS</i>	forward	CATGTCTGGTAACGGCAATG
<i>PBGD/HMBS</i>	reverse	GTACGAGGCTTTCAATGTTG
Spliced <i>XBP1</i>	forward	CGCTTGGGGATGGATGCCCTG
Spliced <i>XBP1</i>	reverse	CCTGCACCTGCTGCGGACT
Total <i>XBP1</i>	forward	GGCATCCTGGCTTGCCTCCA
Total <i>XBP1</i>	reverse	GCCCCCTCAGCAGGTGTTCC

# Human embryonic–stem–cell–derived cardiomyocytes regenerate non–human primate hearts

James J. H. Chong<sup>1,2,3,4,5,†</sup>, Xiulan Yang<sup>1,2,5</sup>, Creighton W. Don<sup>6</sup>, Elina Minami<sup>1,2,5,6</sup>, Yen-Wen Liu<sup>1,2,5</sup>, Jill J. Weyers<sup>1,2,5</sup>, William M. Mahoney Jr<sup>1,2,5</sup>, Benjamin Van Biber<sup>1,2,5</sup>, Savannah M. Cook<sup>7</sup>, Nathan J. Palpant<sup>1,2,5</sup>, Jay A. Gantz<sup>1,2,5,8</sup>, James A. Fugate<sup>1,2,5</sup>, Veronica Muskheli<sup>1,2,5</sup>, G. Michael Gough<sup>9</sup>, Keith W. Vogel<sup>9</sup>, Cliff A. Astley<sup>9</sup>, Charlotte E. Hotchkiss<sup>9</sup>, Audrey Baldessari<sup>9</sup>, Lil Pabon<sup>1,2,5</sup>, Hans Reinecke<sup>1,2,5</sup>, Edward A. Gill<sup>6</sup>, Veronica Nelson<sup>10</sup>, Hans-Peter Kiem<sup>5,10</sup>, Michael A. Laflamme<sup>1,2,5</sup> & Charles E. Murry<sup>1,2,5,6,8</sup>

**Pluripotent stem cells provide a potential solution to current epidemic rates of heart failure<sup>1</sup> by providing human cardiomyocytes to support heart regeneration<sup>2</sup>. Studies of human embryonic–stem–cell–derived cardiomyocytes (hESC–CMs) in small–animal models have shown favourable effects of this treatment<sup>3–7</sup>. However, it remains unknown whether clinical–scale hESC–CM transplantation is feasible, safe or can provide sufficient myocardial regeneration. Here we show that hESC–CMs can be produced at a clinical scale (more than one billion cells per batch) and cryopreserved with good viability. Using a non–human primate model of myocardial ischaemia followed by reperfusion, we show that cryopreservation and intra–myocardial delivery of one billion hESC–CMs generates extensive remuscularization of the infarcted heart. The hESC–CMs showed progressive but incomplete maturation over a 3–month period. Grafts were perfused by host vasculature, and electromechanical junctions between graft and host myocytes were present within 2 weeks of engraftment. Importantly, grafts showed regular calcium transients that were synchronized to the host electrocardiogram, indicating electromechanical coupling. In contrast to small–animal models<sup>7</sup>, non–fatal ventricular arrhythmias were observed in hESC–CM–engrafted primates. Thus, hESC–CMs can remuscularize substantial amounts of the infarcted monkey heart. Comparable remuscularization of a human heart should be possible, but potential arrhythmic complications need to be overcome.**

Human pluripotent stem cells have indisputable cardiomyocyte–generating abilities and have been extensively investigated for repair of the injured heart<sup>3,4,6–10</sup>. These stem cells are derived either from developing blastocysts (human embryonic stem (ES) cells) or from reprogrammed somatic cells (induced pluripotent stem cells (iPSCs))<sup>11</sup>. Although iPSCs have promising therapeutic potential<sup>12</sup>, a number of factors are likely to slow their regulatory approval<sup>2</sup>. Human ES–cell derivatives, on the other hand, are already being tested in humans for retinal diseases and spinal cord injury<sup>13,14</sup>. These indications require small numbers of differentiated cells, ranging from  $10^4$  to  $10^7$ . By contrast, cardiac repair will require orders of magnitude more cells, because a billion cardiomyocytes are lost after a typical infarct<sup>2</sup>. At present it is unknown whether this large–scale production of hESC–CMs is feasible. Furthermore, it remains unclear whether the favourable cardiac repair findings in small–animal models will be reproduced in more clinically relevant large–animal models. As an important translational step towards creating a viable clinical therapy, we investigated the ability of exogenously delivered hESC–CMs to engraft and electrically couple to host myocardium in a non–human primate model of myocardial infarction.

Notably, this model provides a heart size and rate more comparable to the human.

Extrapolating results from our previous studies in smaller mammals, where  $10^6$  cardiomyocytes were required in mice,  $10^7$  in rats and  $10^8$  in guinea pigs<sup>6–8,15</sup>, we reasoned that sufficient engraftment in the larger non–human primate heart required delivery of  $1 \times 10^9$  cells. Feasibility of this large–scale hESC–CM delivery requires cryopreservation of cells, which we validated in an established immunodeficient mouse model of myocardial infarction<sup>15</sup>. Similar to previous reports<sup>16</sup>, we found no adverse impact of cryopreservation on hESC–CM graft size (Extended Data Fig. 1). Therefore, delivery of cryopreserved hESC–CMs seems to be a sound strategy for large–scale transplantation in large animals or humans.

We previously used zinc–finger nuclease (ZFN)–mediated gene targeting to create hESC–CMs (H7 parental ES–cell line) stably expressing the genetically encoded fluorescent calcium indicator GCaMP3 from the *AAVS1* locus<sup>7</sup> (Extended Data Fig. 2a). These were used to prove exogenously delivered hESC–CMs could electrically couple to the host heart in a guinea pig model of myocardial infarction<sup>7</sup>. For the first two non–human primate experiments we used this same cell line. Routine karyotyping after two experiments revealed duplication of the long arm of chromosome 20 (Extended Data Fig. 3a). Reanalysis of two previous karyotypes from this line revealed this subtle duplication to be present in cells delivered to both monkeys. As the effect of this abnormality on hESC–CM engraftment and function is unknown, we created another karyotypically normal GCaMP3 human ES–cell line for comparison. The ZFN approach was again used to target the GCaMP3 construct to the *AAVS1* locus (Extended Data Fig. 2a) in Rockefeller University embryonic stem cell line 2 (RUES2) human ES cells. Southern blotting revealed correct targeting of the construct (Extended Data Fig. 2b) and karyotyping was normal after expansion (Extended Data Fig. 3b). For both of these GCaMP3 ES–cell lines we used our well–established monolayer protocol of directed differentiation (as described earlier) to produce a high yield of cardiomyocytes<sup>8</sup>. Flow cytometry was used to assess cardiomyocyte purity, and the hESC–CMs used in these studies were  $73 \pm 12\%$  positive for cardiac troponin T (cTnT; Extended Data Fig. 4). Spontaneous beating was observed *in vitro* for hESC–GCaMP3–CMs with robust fluorescence with each contractile cycle (Supplementary Videos 1 and 2).

Seven pigtail macaques (*Macaca nemestrina*) were used for the study without randomization (Table 1). Myocardial infarction was created by ischaemia followed by reperfusion using a percutaneous balloon catheter 2 weeks before hESC–CM delivery, with immunosuppression

<sup>1</sup>Center for Cardiovascular Biology, University of Washington, Seattle, Washington 98109, USA. <sup>2</sup>Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington 98109, USA. <sup>3</sup>Department of Cardiology Westmead Hospital, Westmead, New South Wales 2145, Australia. <sup>4</sup>School of Medicine, University of Sydney, Sydney, New South Wales 2006, Australia. <sup>5</sup>Department of Pathology, University of Washington, Seattle, Washington 98195, USA. <sup>6</sup>Department of Medicine/Cardiology, University of Washington, Seattle, Washington 98195, USA. <sup>7</sup>Department of Comparative Medicine, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington 98109, USA. <sup>8</sup>Department of Bioengineering, University of Washington, Seattle, Washington 98195, USA. <sup>9</sup>Washington National Primate Research Center, University of Washington, Seattle, Washington 98195, USA. <sup>10</sup>Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. <sup>†</sup>Present address: University of Sydney School of Medicine, Sydney, New South Wales 2006, Australia and Westmead Millennium Institute and Westmead Hospital, Westmead, New South Wales 2145, Australia.



**Table 1 | Macaque characteristics with morphometry and calcium imaging summary**

Animal identifier	Sex	Age	Body weight (kg)	Heart weight (g)	LV weight (g)	Treatment	Endpoint	Infarct mass (g)	Infarct size (% LV)	Graft mass (g)	Graft size (%LV)	Graft coupled (%)
P2	F	10y 6m	8.6	39	23.3	No-cell control	2 weeks (sham)	1.7	7.3	N/A	N/A	N/A
P3	F	11y 8m	9.2	38	14.9	H7-GCAMP3- CM	2 weeks (cells)	0.8	5.3	0.2	1.3	100
P4	M	9y 5m	9.5	37	19.9	H7-GCAMP3- CM	4 weeks (cells)	1.9	9.5	0.2	0.7	100
P5	M	6y	12.3	52	20.7	RUES2-GCAMP3-CM	4 weeks (cells)	0.5	2.5	1.1	5.3	100
P6	M	5y	9.7	48	29.3	RUES2-GCAMP3 -CM	12 weeks (cells)	1.1	3.7	0.3	1.0	100
P7	F	14y	8.4	36	19.5	No-cell control	4 weeks (sham)	2.0	10.4	N/A	N/A	N/A

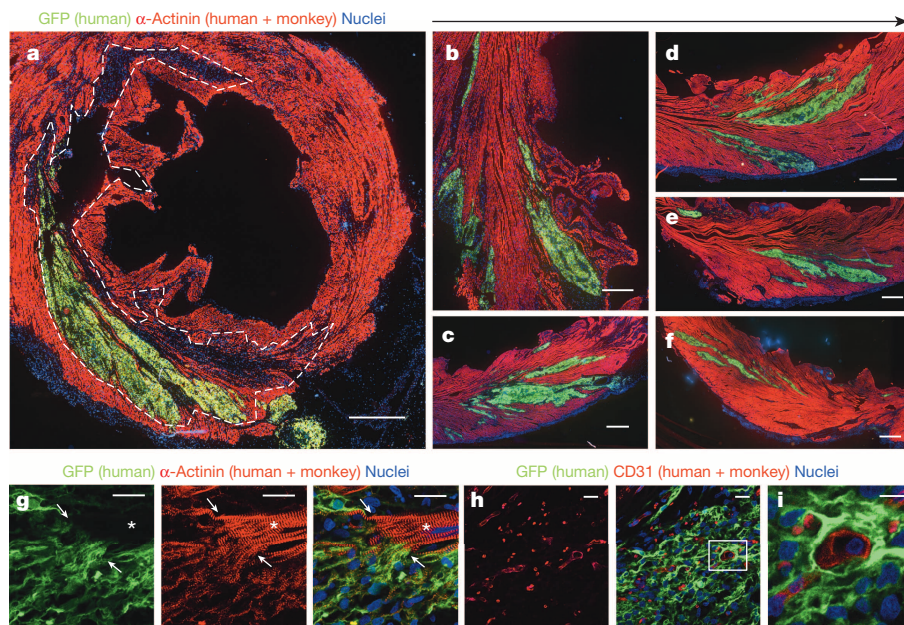
LV, left ventricle; N/A, not applicable.

starting 5 days before cell delivery (see Methods and Extended Data Fig. 5). hESC-CMs were delivered into the infarct region and surrounding border zones under direct surgical visualization using a method optimized to aid cell retention (Extended Data Fig. 6). All macaques underwent full necropsy after being euthanized. Consistent with our previous results<sup>6–8,15</sup>, no macroscopic or microscopic evidence of teratoma or other tumour was detected, and human cells were not identified outside the heart. All macaques had patchy transmural myocardial infarctions. Infarct sizes in sham-treated hearts (at 14 and 28 days after engraftment) were 7.3 and 10.4% of the left ventricle (Table 1), whereas infarcts in cell-treated hearts ranged from 3.7–9.5% of the left ventricle (mean of  $5.2 \pm 1.5\%$ ; Table 1). All hESC-CM-treated monkeys showed extensive remuscularization of the infarct areas (Fig. 1a–g and Extended Data Fig. 7). Graft size, calculated on the basis of green fluorescence protein (GFP) expression, ranged from 0.7–5.3% of the left ventricle (mean of  $2.1 \pm 1.1\%$ ; Table 1), averaging 40% of the infarct volume. Greater than 98% of engrafted human cells expressed the sarcomeric protein  $\alpha$ -actinin (Extended Data Fig. 8a), indicating that almost all graft cells were cardiomyocytes. Furthermore, these hESC-CMs showed increased maturation from day 14 to day 84, as evidenced by increased myofibril alignment, sarcomere registration and cardiomyocyte diameter (Fig. 2a–c and Extended Data Fig. 8b–f). As these conclusions are drawn from small animal numbers per time point ( $n = 1$  each for day 14 and day 84,  $n = 2$  for day 28), maturation will require further validation. The cardiomyocyte diameter of day 84 grafts was  $10.9 \pm 2 \mu\text{m}$ , approximately the size of normal adult monkey cardiomyocytes ( $10.1 \mu\text{m}$ ) and approaching the 11–13  $\mu\text{m}$  diameter seen in normal adult human hearts<sup>17</sup>. Additionally, a maturation gradient was apparent, with cardiomyocytes at the edge of grafts exhibiting greater maturation than those within the central core (Fig. 2f–k). There were frequent host–graft contacts (Fig. 1g) where nascent intercalated disks formed and expressed

the adherens junction protein N-cadherin and the gap junction protein connexin 43. From day 14 to day 84 the expression of these junctional proteins increased substantially (Fig. 2d, e, l–q). Few CD3<sup>+</sup> T lymphocytes or CD20<sup>+</sup> B lymphocytes were found within or around the hESC-CM grafts, suggesting that our immunosuppression successfully prevented graft rejection (Extended Data Fig. 9).

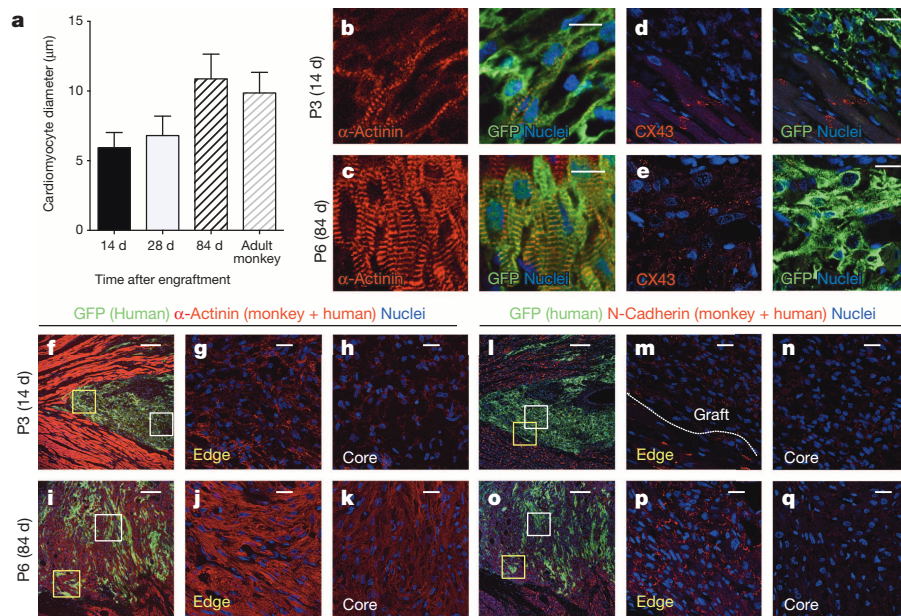
hESC-CM grafts were perfused by host vessels, as evidenced by anti-CD31 immunostaining without GFP co-expression (Fig. 1h, i). Microcomputed tomography was used to image the three-dimensional structure of the coronary vasculature, which was correlated to aligned histological sections, permitting analysis of coronary anatomy within the graft, scar and remote myocardium (Fig. 3a–d and Supplementary Video 3). Graft and scar regions were integrated into the three-dimensional vascular network, revealing arteries and veins supplying the hESC-CM graft that were connected to the host system. This shows, to our knowledge for the first time, that large hESC-CM grafts are successfully perfused by host vasculature and are viable long term.

To investigate electromechanical coupling of hESC-CM grafts to the host, hearts from all macaques were subjected to *ex vivo* fluorescent imaging using a modified Langendorff perfusion system (Supplementary Video 4). Hearts were perfused with 2,3-butanedione monoxime (BDM, a myosin crossbridge inhibitor) to uncouple electrical cardiomyocyte excitation from mechanical contraction. This removed confounding motion artefacts and prevented indirect graft activation by passive stretching. Epicardial fluorescent calcium transients were seen in all hESC-CM-treated hearts, indicating electrical activation of the cardiomyocyte grafts (Fig. 4a–d and Supplementary Videos 5, 6). Furthermore, 100% of the visible hESC-CM grafts in every monkey showed electromechanical coupling to the host heart (Table 1). Graft–host coupling was evidenced by epicardial fluorescent transients that were synchronous with the host electrocardiogram (ECG) QRS complexes during



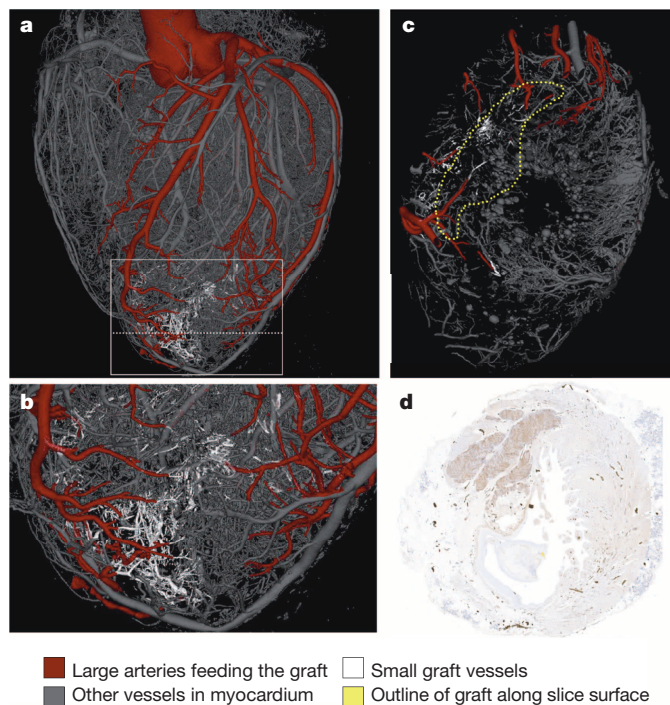
**Figure 1 | Remuscularization of the infarcted macaque heart with human cardiomyocytes.** a–i, Confocal immunofluorescence of macaque hearts subjected to myocardial infarction and transplantation of hESC-CMs. Grafts were studied at day 14 (a–g) and day 84 post-engraftment (h–i). a, Remuscularization of a substantial portion of the infarct region (dashed line) with hESC-CMs co-expressing GFP. The contractile protein  $\alpha$ -actinin (red) is expressed by both monkey and human cardiomyocytes. Scale bar, 2,000  $\mu\text{m}$ . b–f, Images from the peri-infarct region of the same heart shown in a, demonstrating extensive hESC-CM engraftment. Scale bars: 1,000  $\mu\text{m}$  (b–e); 200  $\mu\text{m}$  (f). g, Graft–host interface (arrows) at day 14 with interconnected  $\alpha$ -actinin- (red) expressing cardiomyocytes (arrows). Note that host sarcomeric cross-striations (asterisks) show greater alignment than hESC-CM graft. Scale bars, 25  $\mu\text{m}$ . h–i, Day 84 hESC-CM grafts contain host-derived blood vessels lined by CD31<sup>+</sup> endothelial cells. Scale bars, 20  $\mu\text{m}$ . Inset scale bar, 10  $\mu\text{m}$ .





**Figure 2 | Human cardiomyocyte grafts mature with time from engraftment.** **a**, Cardiomyocyte diameter of hESC-CMs shows significant increase from 14 ( $n = 1$ ) to 28 ( $n = 2$ ) days and from 28 to 84 ( $n = 1$ ) days after engraftment. Adult monkeys ( $n = 2$ ). From each animal 200–400 cells were counted from three histological sections at varied left ventricular levels. Mean  $\pm$  standard error of the mean (s.e.m.) is shown. **b–q**, Confocal immunofluorescence of macaque hearts subjected to myocardial infarction and transplantation of hESC-CMs 14 days (**b, d, f–h, l–n**) or 84 days (**c, e, i–k, o–q**) after engraftment. P3 and P6 are animal identifiers. Increased myofibril content, sarcomere alignment and cardiomyocyte size in hESC-CMs (GFP<sup>+</sup>) are seen in longer-term grafts (**b, c**). Connexin43 (CX43) expression is not evident in hESC-CM grafts at 14 days but is seen at 84 days (**d, e**). Cardiomyocytes at the edges of grafts (**j, k, m, p**) show greater maturation compared with those at the central core (**h, k, n, q**), as evidenced by increased size,  $\alpha$ -actinin staining intensity, sarcomere alignment (**g–k**) and N-cadherin expression (**m–q**). Scale bars for panels **f, i, l** and **o**, 100  $\mu$ m. All other scale bars, 20  $\mu$ m. Yellow and white boxes correspond to higher-power fields of graft edge and core, respectively.

spontaneous depolarization (Fig. 4e). hESC-CM grafts retained 1:1 coupling to host myocardium during atrial pacing at rates of up to 240 beats per minute, the highest rate tested (Fig. 4f–h).

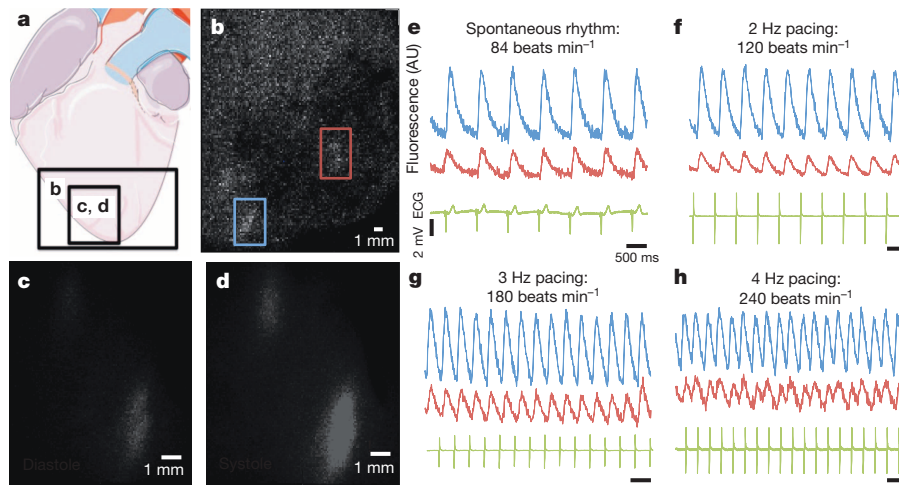


**Figure 3 | Blood vessels extend from the host coronary network into the graft.** **a–c**, Three-dimensional rendered microcomputed tomography of heart perfused with Microfil at 3 months after hESC-CM injection. **b**, Higher-power view of boxed area from **a**. **c**, Cross-sectional cut plane through the heart at the location of the dotted line in **a**. Arteries perfusing the graft are red, other vessels are grey in the uninjured cardiac tissue, or white within the graft. The vessels within the graft are better visualized in Supplementary Video 6. **d**, A histological section of the heart shown in **a–c** was immunostained with an anti-GFP antibody to mark the hESC-CM graft (brown). This section corresponds to the same location of the cross-sectional cut plane in **c**. Black dots are Microfil within coronary vessels.

To explore the electrophysiological consequences of our hESC-CM grafts, we analysed ECGs obtained by telemetry from the time of infarction until death. Continuous ECG recordings were taken regularly and 24 h periods (midnight to midnight) were analysed. Control macaques with myocardial infarctions and sham (vehicle only) injections maintained normal sinus rhythm with heart rates of 100–130 beats per minute throughout the experiment (Fig. 5a). No arrhythmias were noted in hESC-CM-treated monkeys during the period after myocardial infarction but before hESC-CM delivery (Fig. 5e–h). By contrast, all macaques that received hESC-CMs showed arrhythmias. These included premature ventricular contractions and runs of ventricular tachycardia (defined as wide QRS complex ( $>60$  ms) with rate  $>180$  beats  $\text{min}^{-1}$ ; Fig. 5c). Frequent wide QRS complex rhythms with rates similar to baseline (accelerated idioventricular rhythm; Fig. 5b) were also observed. Notably, all animals remained conscious and in no distress during all periods of arrhythmia.

To investigate left ventricular function, we performed transoesophageal echocardiography before myocardial infarction, before hESC-CM delivery and immediately before the end of the experiment (Extended Data Fig. 10b). Multiple trans-oesophageal and deep trans-gastric views were analysed by cardiologists blinded to experimental details. We were unable to obtain images of sufficient quality for analysis from one control animal. The other control demonstrated a decline in ejection fraction after myocardial infarction that was unchanged after sham cell injection. The hearts receiving hESC-CMs showed variable responses, some exhibiting an increased ejection fraction after treatment and others showing no improvement. Owing to the small group sizes, no statistically significant effects were noted.

These experiments demonstrate that hESCs can be grown, differentiated into cardiomyocytes and cryopreserved at a scale sufficient to treat a large-animal model of myocardial infarction. With further refinements in manufacturing, the scale up to trials in human patients seems feasible. Large-animal models are important forerunners to human trials, because they impart real-world rigour to issues such as cell production, delivery and end-point analyses, while permitting mechanistic studies not possible in patients<sup>18,19</sup>. We observed extensive remuscularization of the infarcts in all animals, with grafts averaging 40% of infarct mass. Importantly, all of the human cardiomyocytes showed complete electrical coupling to the primate heart and responded normally to pacing up to 240 beats per minute (the fastest rate attempted).



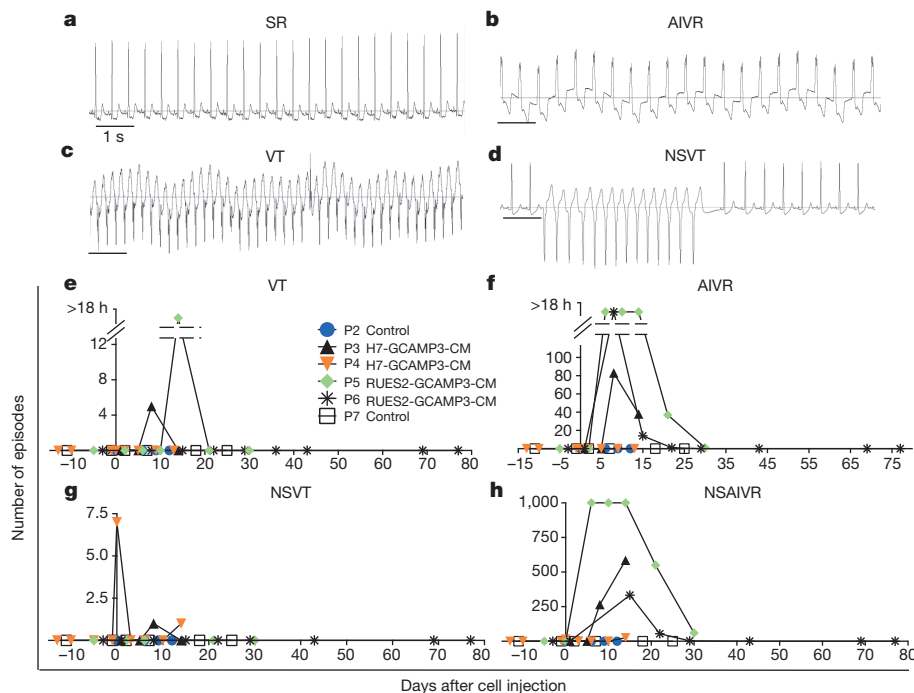
**Figure 4 | Human cardiomyocytes are electrically coupled 1:1 to the infarcted host macaque heart after transplantation.** **a**, Diagram showing regions of the infarcted macaque heart visualized in **b–d**. Analysis shown is from *ex vivo* imaging 14 days after hESC-CM delivery. **b**, Still image from low-power fluorescence video showing regions of hESC-CM engraftment (red and blue rectangles). **c**, **d**, Still images of calcium indicator GCaMP3-positive hESC-CM grafts (bottom left of panel **b**) during diastole (**c**) and systole (**d**).

Note the gain of fluorescence during systole. **e–h**, GCaMP3 fluorescence intensity (arbitrary units (AU)) and ECG versus time for the grafted regions of interest shown in **b**. **e–h**, Each graft region shows 1:1 coupling synchronous with host ventricular contraction (ECG QRS complex) during spontaneous rhythm (**e**) or atrial pacing (**f–h**). All hESC-CM grafts identified in every transplanted animal showed 1:1 coupling.

The coupling seen in this study was greater than that observed in our guinea pig model, where only 60% of recipient hearts had grafts that were synchronized with the host<sup>7</sup>. This enhanced coupling may have resulted from the use of an ischaemia–reperfusion model, which gives patchier infarcts with more peninsulas of viable host tissue than the guinea pig cryo-injury model.

Our previous studies in mice<sup>15</sup>, rats<sup>6,8</sup> and guinea pigs<sup>7</sup> gave no evidence of arrhythmias after hESC-CM engraftment, whereas here we consistently observed arrhythmias. There are several possible mechanisms

for the observed arrhythmias, including re-entrant circuits or graft automaticity<sup>20–22</sup>. Further studies are required to distinguish between these possibilities. The most likely reasons why arrhythmias were observed in monkeys but not in smaller animals seem to be differences in heart size and rate. Regarding size, the larger hearts of adult macaques (37–52 g) compared with the hearts of mice (0.15 g), rats (1 g) and guinea pigs (3 g) allows for more hESC-CMs to be delivered, and the resultant grafts are approximately tenfold larger than the largest obtained in other species<sup>7</sup>. Ventricular depolarization over integrated but relatively



**Figure 5 | Ventricular arrhythmias after hESC-CM transplantation.** **a–d**, Representative traces from macaque telemetric ECG recordings showing normal sinus rhythm (SR; **a**), accelerated idioventricular rhythm (AIVR; **b**), ventricular tachycardia (VT; **c**) and non-sustained VT (NSVT; **d**). Scale bar, 1 s. **e–h**, Frequency of arrhythmias is highest within the first 2 weeks after

hESC-CM transplantation. P2–7 designations are animal identifiers. Animals receiving vehicle only (no cells, P2 and P7) remained in SR throughout. Interrupted y-axis in **e**, **f** denotes reduced number of episodes but increased total duration of arrhythmias (VT or AIVR for more than 18 h per 24 h period). NSAIVR, non-sustained accelerated idioventricular rhythm.



immature hESC-CM grafts may slow conduction of the overall wave front. Although not problematic over short distances in small grafts, over longer distances (in large grafts) this may favour formation of re-entrant loops. It is noteworthy that the animal (P5) with the largest hESC-CM graft size also had the highest frequency of arrhythmia. Another important factor is the species-specific heart rates (macaques 100–130 beats min<sup>-1</sup> versus guinea pigs 230 beats min<sup>-1</sup>, rats ~400 beats min<sup>-1</sup> and mice ~600 beats min<sup>-1</sup>). Faster spontaneous rates will favour ventricular capture from native conduction pathways rather than graft automaticity or re-entrant loops, and this would probably prevent sustained ventricular arrhythmias. These factors are relevant to clinical translation given that the human heart is larger (300 g) with a slower basal rate (70 beats min<sup>-1</sup>) than that of macaques.

The principal limitations of this study are the small numbers of animals used and their relatively small infarct sizes. Both limitations stem from the high cost and value of the primate model. Consequently, we cannot determine with statistical certainty that the observed arrhythmias directly result from transplanted hESC-CMs. Larger studies will be required to assess this and the treatment effects on cardiac function. Importantly, infarct sizes in this study were smaller than the clinically severe infarcts that might benefit most from hESC-CM therapy. Larger infarcts, in human hearts, might manifest more arrhythmias. Because ventricular arrhythmias can be life threatening, they need to be understood mechanistically and managed en route to safe clinical translation. Nevertheless, the extent of remuscularization and electromechanical coupling seen here encourages further development of human cardiomyocyte transplantation as a clinical therapy for heart failure.

## METHODS SUMMARY

Human ES cells were differentiated into cardiomyocytes by induction with activin A and BMP4, as previously reported<sup>7,8</sup>. To enhance engraftment cardiomyocytes were subjected to heat shock followed by treatment with a pro-survival cocktail before cryopreservation. GCaMP3-positive human ES cells were generated by ZFN-mediated targeting to the AAVS1 locus, following methods described previously<sup>7</sup>. Details of mouse and macaque procedures are provided in Methods. Microcomputed tomography was performed as previously described<sup>23</sup> with minor modifications. All procedures complied with the regulations of and were approved by the University of Washington Institutional Animal Care and Use Committee.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 September 2013; accepted 6 March 2014.**

**Published online 30 April; corrected online 11 June 2014 (see full-text HTML version for details).**

- Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2095–2128 (2012).
- Laflamme, M. A. & Murry, C. E. Heart regeneration. *Nature* **473**, 326–335 (2011).
- Caspi, O. *et al.* Transplantation of human embryonic stem cell-derived cardiomyocytes improves myocardial performance in infarcted rat hearts. *J. Am. Coll. Cardiol.* **50**, 1884–1893 (2007).
- van Laake, L. W. *et al.* Human embryonic stem cell-derived cardiomyocytes survive and mature in the mouse heart and transiently improve function after myocardial infarction. *Stem Cell Res. (Amst.)* **1**, 9–24 (2007).
- Laflamme, M. A. *et al.* Formation of human myocardium in the rat heart from human embryonic stem cells. *Am. J. Pathol.* **167**, 663–671 (2005).
- Fernandes, S. *et al.* Human embryonic stem cell-derived cardiomyocytes engraft but do not alter cardiac remodeling after chronic infarction in rats. *J. Mol. Cell. Cardiol.* **49**, 941–949 (2010).
- Shiba, Y. *et al.* Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. *Nature* **489**, 322–325 (2012).

- Laflamme, M. A. *et al.* Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nature Biotechnol.* **25**, 1015–1024 (2007).
- Blin, G. *et al.* A purified population of multipotent cardiovascular progenitors derived from primate pluripotent stem cells engrafts in postmyocardial infarcted nonhuman primates. *J. Clin. Invest.* **120**, 1125–1139 (2010).
- Bel, A. *et al.* Composite cell sheets: a further step toward safe and effective myocardial regeneration by cardiac progenitors derived from embryonic stem cells. *Circulation* **122**, S118–S123 (2010).
- Yang, X., Pabon, L. & Murry, C. E. Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ. Res.* **114**, 511–523 (2014).
- Mauritz, C. *et al.* Induced pluripotent stem cell (iPSC)-derived Flk-1 progenitor cells engraft, differentiate, and improve heart function in a mouse model of acute myocardial infarction. *Eur. Heart J.* **32**, 2634–2641 (2011).
- Schwartz, S. D. *et al.* Embryonic stem cell trials for macular degeneration: a preliminary report. *Lancet* **379**, 713–720 (2012).
- Bretzner, F., Gilbert, F., Baylis, F. & Brownstone, R. M. Target populations for first-in-human embryonic stem cell research in spinal cord injury. *Cell Stem Cell* **8**, 468–475 (2011).
- Robey, T. E., Saiget, M. K., Reinecke, H. & Murry, C. E. Systems approaches to preventing transplanted cell death in cardiac repair. *J. Mol. Cell. Cardiol.* **45**, 567–581 (2008).
- Xu, C. *et al.* Efficient generation and cryopreservation of cardiomyocytes derived from human embryonic stem cells. *Regen. Med.* **6**, 53–66 (2011).
- Hoshino, T., Fujiwara, H., Kawai, C. & Hamashima, Y. Myocardial fiber diameter and regional distribution in the ventricular wall of normal adult hearts, hypertensive hearts and hearts with hypertrophic cardiomyopathy. *Circulation* **67**, 1109–1116 (1983).
- Gandolfi, F. *et al.* Large animal models for cardiac stem cell therapies. *Theriogenology* **75**, 1416–1425 (2011).
- van der Spoel, T. I. *et al.* Human relevance of pre-clinical studies in stem cell therapy: systematic review and meta-analysis of large animal models of ischaemic heart disease. *Cardiovasc. Res.* **91**, 649–658 (2011).
- Kehat, I. *et al.* Electromechanical integration of cardiomyocytes derived from human embryonic stem cells. *Nature Biotechnol.* **22**, 1282–1289 (2004).
- Chen, H. S., Kim, C. & Mercola, M. Electrophysiological challenges of cell-based myocardial repair. *Circulation* **120**, 2496–2508 (2009).
- Dixon, J. A. & Spinale, F. G. Large animal models of heart failure: a critical link in the translation of basic science to clinical practice. *Circ Heart Fail* **2**, 262–271 (2009).
- Weyers, J. J. *et al.* Effects of cell grafting on coronary remodeling after myocardial infarction. *J. Am. Heart Assoc.* **2**, e000202 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Dupras, B. Brown, D. Rocha, E. Wilson, C. English, J. Randolph-Habecker and T. Goodpaster for assistance with these experiments. This work was supported by National Institutes of Health grants P01HL094374, R01HL084642, U01HL100405 and P01GM081619 and an Institute of Translational Health Sciences/Primate Center Ignition Award. J.J.H.C. was supported by National Health and Medical Research Council of Australia Overseas Training and Australian-American Fulbright Commission Fellowships. X.Y. is supported by an American Heart Association post-doctoral scholarship 12POST11940060. J.J.W. is supported by an American Heart Association post-doctoral scholarship 12POST9330030. H.-P.K. is a Markey Molecular Medicine investigator and the recipient of the Jose Carreras/E.D. Thomas Chair for Cancer Research.

**Author Contributions** J.J.H.C., X.Y., C.W.D., E.M., L.P., H.R., H.-P.K., M.A.L. and C.E.M. designed the study. J.J.H.C. and E.M. performed mouse transplantation experiments. J.J.H.C. developed telemetry and analysed recordings. J.J.H.C., C.W.D., C.E.M., G.M.G., K.W.V., C.A.A., E.M. and V.N. performed macaque surgery and procedures. J.J.H.C., E.M., E.A.G. and C.E.H. performed echocardiography and E.M., E.G. and Y.-W.L. performed analysis. A.B. performed necropsies and non-cardiac histopathology. GCaMP3 visualization experiments were carried out and analysed by X.Y. and J.J.H.C. GCaMP3-expressing human ES cells were created by N.J.P., J.A.G. and B.V.B. hESC-CM production was by J.J.H.C., B.V.B., S.M.C., J.A.F. and M.A.L. Microcomputed tomography experiments were performed by J.J.W. and W.M.M. Jr. Immunohistochemistry was performed and analysed by J.J.H.C., V.M. and Y.-W.L. Figures were created by J.J.H.C. with assistance from X.Y., J.J.W., Y.-W.L., N.J.P. and V.M. The manuscript was written principally by J.J.H.C. and C.E.M.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.E.M. ([murry@uw.edu](mailto:murry@uw.edu)).



## METHODS

**Cell preparation.** Undifferentiated H7 (ref. 11) or RUES2 human ES cells<sup>24</sup> were expanded using mouse embryonic fibroblast-conditioned medium (MEF-CM)<sup>25</sup> supplemented with basic fibroblast growth factor (R&D Systems). The H7 line was obtained from the WiCell Research Institute and the RUES2 line from Rockefeller University. Both lines were regularly karyotyped and tested for mycoplasma. Human ES cells were then differentiated into cardiomyocytes using a previously reported directed differentiation protocol. Briefly, activin A (R&D Systems) and bone morphogenetic protein 4 (BMP4, R&D) are applied to defined, serum-free, monolayer culture conditions<sup>8,26</sup>. hESC-CMs were collected and cryopreserved after 16–20 days of CM differentiation. One day before collection, cells were subjected to a pro-survival 'cocktail' (PSC) protocol, previously shown to enhance engraftment after transplantation<sup>8</sup>. Briefly, cultures were heat-shocked with a 30 min exposure to 43 °C medium, followed by RPMI-B27 medium supplemented with IGF1 (100 ng ml<sup>-1</sup>, Peprotech) and cyclosporine A (0.2 mM, Sandimmune, Novartis). One day later, cultures were collected with 0.25% trypsin per 0.5 mM EDTA (Invitrogen) and cryopreserved as described previously<sup>7</sup>. Immediately before transplantation, cells were thawed at 37 °C, washed with RPMI, and suspended in 1.5 ml volume (per animal) of modified PSC consisting of 50% (v/v) growth-factor-reduced Matrigel, supplemented with BCL-xl BH4 (cell-permeant TAT peptide, 50 nM, Calbiochem), cyclosporine A (200 nM, Wako), IGF1 (100 ng ml<sup>-1</sup>, Peprotech) and pinacidil (50 mM, Sigma).

**Generation of the GCaMP-reporter human ES-cell line.** A transgene encoding for the constitutive expression of GCaMP3 was inserted into the AAVS1 locus in H7 and RUES2 human ES cells, using methods adapted from a previous study<sup>27</sup> (see Extended Data Fig. 2). In brief, the right and left arms of an AAVS1-specific ZFN were *de novo* synthesized (Genscript) and cloned into a single polycistronic plasmid in which the expression of each was driven by an independent human PGK promoter. A second polycistronic vector was generated in which (approximately 800 bp) homology arms flanking the AAVS1 ZFN cut site (pZDonor, Sigma Aldrich) surrounded a 5.1 kb insert with two elements: a cassette in which the CAG promoter drives expression of GCaMP3 (Addgene, plasmid #22692) and a second cassette encoding for PGK-driven expression of neomycin resistance. AAVS1 ZFN (5 mg) and AAVS1 CAG GCaMP3 targeting vector plasmids were co-electroporated (Lonza, Nucleofection system) into human ES cells cultured in MEF-CM supplemented with 10 mM Y-27632. Green fluorescent colonies were isolated and expanded and selected with 40–100 µg ml<sup>-1</sup> G418 (Invitrogen) for 5–10 days.

**Southern blot analysis.** Wild-type and transgenic GCaMP3-positive human ES-cell genomic DNA were digested with the restriction enzymes NdeI and NheI, run on 1% polyacrylamide gel and transferred to a membrane (BioRad Zeta Probe). The membrane was washed in 2× SSC and dried at 80 °C in a hybridization oven for 2 h, followed by 1 h of pre-hybridization in 50% formamide, 0.12 M NaH<sub>2</sub>PO<sub>4</sub>, 0.25 M NaCl, 7% SDS and 1 mM EDTA at 43 °C. A genomic probe was generated using the following primers: GGAGGTGGTGGCTTCTTGG (forward), CGC ATCCCTCCAGAAAGAC (reverse), and neomycin cassette probe: ATGGGA TCGGCCATTGAACAAG (forward), GAAGAAGCTCGTCAAGAAGGCG (reverse). The probes were labelled with p32 dCTP (Amersham Megaprime DNA labelling system) and hybridized overnight in hybridization buffer at 43 °C. After 24 h, the membrane was washed for 20 min with 2× SSC/0.1% SDS followed by 20 min in 0.1× SSC/0.1% SDS. The membrane was then exposed to autoradiographic film for 3 days.

**Animal models.** All procedures complied with and were approved by the University of Washington Animal Care and Use Committee.

**Mouse surgery.** Male, SCID-BEIGE mice of 8 weeks age, (Taconic Farm) were anaesthetized with Avertin, intubated and ventilated before undergoing thoracotomy and ligation of the left anterior descending artery. Immediately after ligation 1 × 10<sup>5</sup> hESC-CMs (freshly isolated or cryopreserved, allocated in a non-randomized and unblinded manner) in a volume of 5 µl was injected directly into the infarct region and surrounding border zones. Five days after myocardial infarction creation mice were euthanized and hearts collected for analysis for detection of human cell grafts by previously described methods<sup>15</sup> (see later and Extended Data Fig. 1).

**Non-human primate surgery.** *M. nemestrina* (8.6–12.3 kg, Washington National Primate Center) of either sex were used for these experiments. Ages are specified in Table 1. Macaques first underwent a 2-week period of acclimation and training to wear a mesh jacket to prevent removal of intravenous (i.v.) catheter. Five days before myocardial infarction, macaques were treated with amiodarone 100 mg daily with feed 5 days before myocardial infarction and continued for a further 10 days after myocardial infarction. For all major surgery macaques were anaesthetized with ketamine and propofol, intubated and ventilated using sevoflurane to maintain anaesthesia. Fentanyl and buprenorphine were administered to provide perioperative and postoperative pain relief. Before each major surgery trans-oesophageal echocardiography

was performed using a Phillips HD-11XE ultrasound machine with an S7 2 MHz trans-oesophageal probe.

Before myocardial infarction creation, an i.v. lidocaine bolus 1 mg kg<sup>-1</sup> and infusion 20 µg kg<sup>-1</sup> min<sup>-1</sup> was used to prevent ventricular arrhythmias. Heparin was delivered i.v. to maintain activated clotting times of 250–350 s to prevent thrombosis. Under fluoroscopic guidance a 5F coronary catheter was used to engage the left main coronary artery. A guide wire and angioplasty balloon was passed into the mid-left anterior descending artery and the balloon inflated for 90 min. Myocardial infarction was confirmed by ST segment elevation on ECG and by subsequent serum assays for cardiac troponin and creatine kinase. For telemetric monitor implantation a CTA-D70 (Data Sciences International) transmitter was placed subcutaneously over the abdomen with leads tunnelled subcutaneously in a modified lead II configuration. Immune suppression was achieved by methylprednisolone i.v. 500 mg on the day before hESC-CM delivery then maintenance doses of 0.1–1.5 mg kg<sup>-1</sup> until monkeys were euthanized, cyclosporine to maintain serum trough levels of 200–250 µg l<sup>-1</sup> from 5 days before hESC-CM delivery until macaques were euthanized and Abatacept (CTLA4 immunoglobulin) 12.5 mg kg<sup>-1</sup> on the day before hESC-CM and every 2 weeks thereafter. To prevent opportunistic infections broad-spectrum antibiotics and anti-fungal agents were administered.

On day 14 after myocardial infarction, macaques were anaesthetized and underwent left thoracotomy. The heart was exposed and a pericardial cradle created. The infarct region was directly visualized and hESC-CMs were delivered intramyocardially into the infarct region and adjacent border zones via 15 injections each of 100 µl volume. Needle tips were placed within a preformed mattress suture, and three injections were delivered via the same epicardial puncture, changing the trajectory of the needle for each. Before withdrawal of the needle the mattress suture was closed around the needle tip to facilitate cell retention. For control macaques, an equal volume of PSC-RPMI vehicle was injected in the same manner as for hESC-CM delivery. hESC-CM-treated animals also received epicardial application of 1–3 tissue-engineering constructs where hESC-CMs were seeded in a collagen scaffold. (These tissue engineered constructs did not adhere to the epicardial surface and were not recovered at the end of the experiment.) Euthanasia was induced by i.v. injection of pentobarbital and phenytoin (Beuthanasia-D) followed by supersaturated KCl and Beuthanasia (CIII). Hearts were removed and perfused with University of Wisconsin cardioplegia solution before transportation on ice for calcium imaging experiments. Seven macaques were subjected to myocardial infarction. One was euthanized 2 days post-infarction (no treatment) because of lower limb ischaemia secondary to arterial thrombosis, and this was the only animal excluded from analysis. All others survived to the completion of the experiment. Two macaques (one cell-treated and one vehicle-only control) were euthanized at day 14, three macaques (two cell-treated and one vehicle-only control) were euthanized at day 28, and one cell-treated monkey was euthanized 84 days (3 months) after hESC-CM delivery. Control and cell-treatment groups were allocated in an unblinded and non-randomized manner.

**PCR detection of human ES-cell grafts.** A high-throughput method of human cell detection was used as previously reported<sup>15</sup>. Briefly, hearts from mice engrafted with hESC-CMs were washed, snap frozen in liquid nitrogen and homogenized using a dis-membrator (Braun). Samples were resuspended in 200 µl of RNase/DNase-free water supplemented with proteinase K and Chelex beads. Samples were centrifuged and a 2 µl sample of the DNA-containing supernatant removed for subsequent PCR using Alu-specific primers. Data were compared to standard curves generated with known human DNA quantities.

**Imaging of GCaMP3-expressing grafts.** Intravital imaging of hearts with GCaMP3-positive grafts was performed on days 14, 28 or 84 after hESC-CM transplantation using *ex vivo* preparation. For these experiments, the heart was mounted on a gravity-fed Langendorff apparatus and then perfused at 100 mm Hg with modified Tyrode solution at 37 °C. The epicardial GCaMP3 signal was then recorded before and after supplementation of the perfusate with 2,3-butanedione monoxime (BDM; 20 mM)<sup>28,29</sup>. GCaMP3 signal was visualized using an epifluorescence stereomicroscope (Nikon, SMZ 1000) equipped with an EXFO X-Cite illumination source. GCaMP3 was excited at 450–490 nm and bandpass filtered (500–550 nm) before detection by an electron-multiplying, charge-coupled device camera (Andor iXon 860 EM-CCD) controlled by Andor Solis software. GCaMP3 image acquisition was typically at 80–140 frames per second (f.p.s.). Signals from the charge-coupled device (CCD) camera and the surface ECG were fed through a computer for digital storage and off-line analysis using Andor software and Labchart.

**Echocardiography.** Images were acquired with an HD11-XE (Phillips) with S7 2 MHz trans-oesophageal probe. Trans-oesophageal four-chamber, two-chamber and short axis views were collected together with deep trans-gastric short-axis views. Functional analysis was performed using XCelera (Phillips) software by two independent cardiologists blinded to experimental conditions.

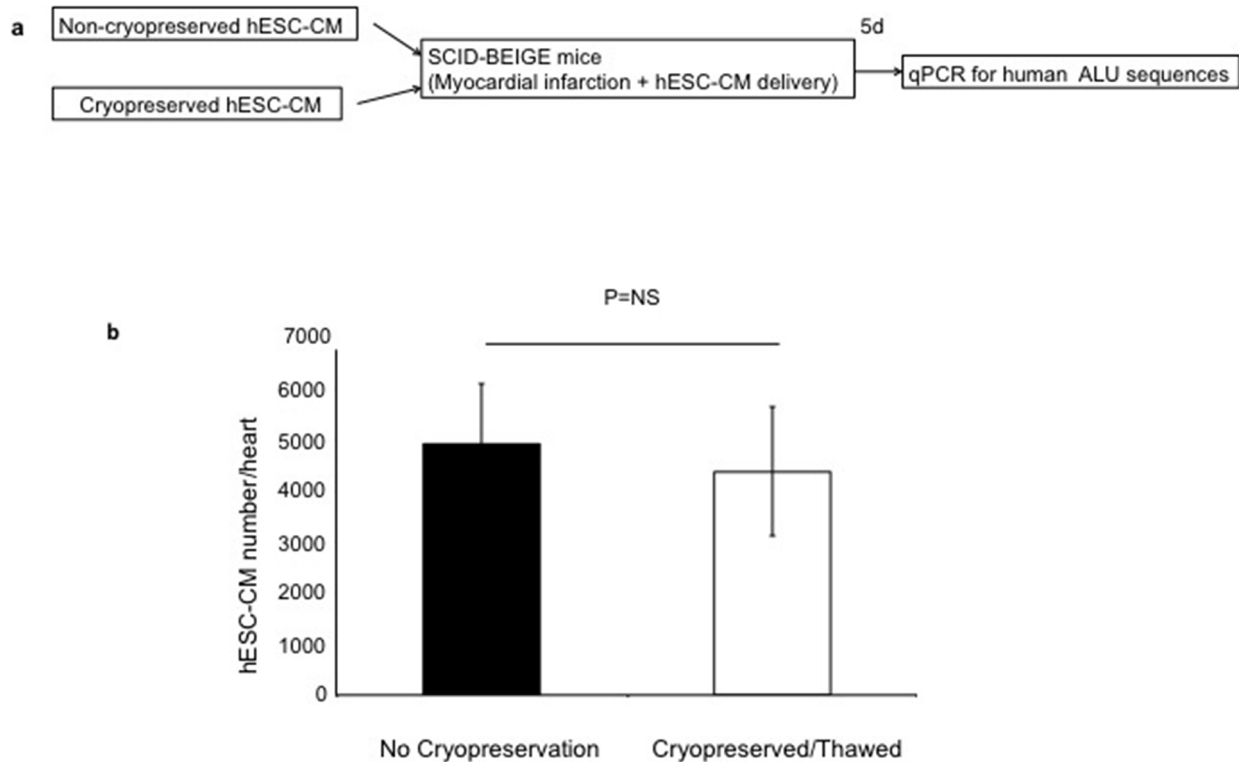
**Telemetric ECG.** ECG recordings were acquired from conscious, freely mobile animals using a Dataquest ART telemetry system (DSI). Recordings from 24 h periods (midnight to midnight) were obtained from macaques with myocardial infarction with or without hESC-CM delivery. All ECG traces were evaluated manually by a cardiologist using Ponemah software (DSI) who determined the total number and frequency of events. Ventricular tachycardia (VT) was defined as a run of four or more premature ventricular complexes (PVCs) with ventricular rate of more than 180 beats per minute. Accelerated idioventricular rhythm (AIVR) was defined as four or more PVCs with a rate of less than 180 beats per minutes. VT or AIVR were considered sustained if the duration was greater than 30 s.

**Histology and immunohistochemistry.** Histological studies were carried out as detailed previously by our group<sup>7,8,30</sup> with some adaptation. For immunohistochemistry, we used the primary antibodies detailed in Extended Data Fig. 10, then either fluorescent secondary antibodies (Alexa-conjugated, species-specific antibodies from Molecular Probes) or the avidin biotin reaction followed by chromogenic detection (ABC kits from VectorLabs). Paraformaldehyde-fixed macaque hearts were dissected to remove the atria and right ventricle before cross-sections were obtained by sectioning parallel to the shortaxis at ~3 mm thickness on a commercial slicer (Berkel). Whole heart, left ventricle and each slice were weighed before tissue processing. For morphometry, infarct regions were identified by Picrosirius red staining and areas calculated using Nanozoomer scanning and software (Hamamatsu). Graft sizes were calculated by anti-GFP staining. All immunofluorescent images were collected by a Nikon A1 Confocal System attached to a Nikon Ti-E inverted microscope platform and using a water-immersion Nikon  $\times 60$  CFI Plan Apo objective lens with 1.2 NA. Image acquisition was performed at room temperature using Nikon NIS Elements 3.1 software to capture 12-bit raw files that were then rescaled to 16-bit images for further processing. All images were collected as a single scan with the pinhole adjusted to 1 Airy unit at  $1,024 \times 1,024$  pixel density. For figure preparation, images were exported into Photoshop CS3 (Adobe). If necessary, brightness and contrast were adjusted for the entire image and the image was cropped. Live cell imaging was performed using a Nikon Eclipse TS100 inverted microscope with white light source and an X-Cite Series 120Q Laser. For calculation of cardiomyocyte diameter, longitudinally sectioned cardiomyocytes were chosen for measurement. Transversely or obliquely cut cardiomyocytes were excluded from morphometric analysis. A point-to-point perpendicular measured line at the position of midnucleus level and the diameter measured using Image J software (version 1.47). At least 200 cardiomyocytes were measured in each animal.

**Microcomputerized tomography scanning and image analysis.** Microcomputerized tomography ( $\mu$ CT) was performed as previously described<sup>23</sup>. Microfilled hearts were imaged in a Skyscan 1076  $\mu$ CT scanner at 35  $\mu$ m spatial resolution using the following settings: 55 kV, 180 mA, 0.5 mm aluminium filter, 220 ms exposure, rotation step of  $0.5^\circ$ ,  $180^\circ$  scan, and  $10\times$  frame averaging. Raw scan data were reconstructed to a three-dimensional slice data set with an isotropic

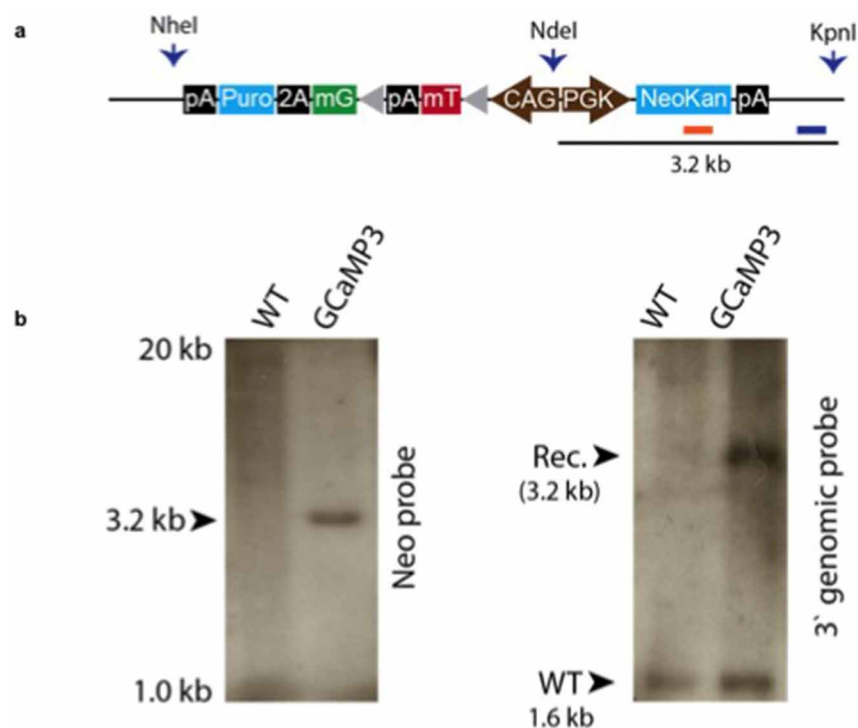
resolution of 35  $\mu$ m using the software NRecon version 1.6.1.0 (Skyscan), and analysed using CTan (Skyscan) and Analyze 10.0 (Mayo Clinic) as follows. Samples were thresholded to a level where vessels separated into distinct entities to allow visualization of individual networks. Non-vascular Microfil (for example, in the atria, aorta, coronary sinus, and so on) was digitally removed in Analyze using the 'Image Segmentation' module. Then, delineation of graft and/or scar tissue was drawn in CTan: histological sections of each heart slice (sliced at 2 mm thickness) stained to highlight the graft and scar (Picrosirius red (scar) and GFP (graft); see later) were imported into the microcomputerized tomography ( $\mu$ CT) three-dimensional data set by aligning and replacing two-dimensional  $\mu$ CT slices at 2 mm intervals. The graft and scar regions were manually outlined on these histological pictures, and the region of interest (ROI) interpolation function in CTan extended the ROI from the manually outlined slices across all slices to produce a three-dimensional representation of the graft or scar (volume of interest (VOI)). The resulting graft/scar VOI was then used to distinguish vessel location (that is, graft, scar or uninjured cardiac tissue) in subsequent analyses. Individual vessel segmentation to determine vessel identity and branching pattern was performed using Analyze with the graft/scar VOI imported from CTan. Arterial/venous identity in three-dimensional renderings was assigned by determining the origin of each vascular network (for example, aorta, coronary sinus, and so on). **Statistical analysis.** All values are expressed as mean  $\pm$  s.e.m. Statistical analyses were performed using Graphpad Prism software, with the threshold for significance level set at  $P < 0.05$ . For murine cryopreservation graft analysis study and india ink injection experiments, paired  $t$ -test analysis of means was used.

24. Gantz, J. A. *et al.* Targeted genomic integration of a selectable floxed dual fluorescence reporter in human embryonic stem cells. *PLoS ONE* **7**, e46971 (2012).
25. Xu, C. *et al.* Feeder-free growth of undifferentiated human embryonic stem cells. *Nature Biotechnol.* **19**, 971–974 (2001).
26. Zhu, W. Z., Van Biber, B. & Laflamme, M. A. Methods for the derivation and use of cardiomyocytes from human pluripotent stem cells. *Methods Mol. Biol.* **767**, 419–431 (2011).
27. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature Biotechnol.* **27**, 851–857 (2009).
28. Biermann, M. *et al.* Differential effects of cytochalasin D and 2,3 butanedione monoxime on isometric twitch force and transmembrane action potential in isolated ventricular muscle: implications for optical measurements of cardiac repolarization. *J. Cardiovasc. Electrophysiol.* **9**, 1348–1377 (1998).
29. Laurita, K. R. & Singal, A. Mapping action potentials and calcium transients simultaneously from the intact heart. *Am. J. Physiol. Heart Circ. Physiol.* **280**, H2053–H2060 (2001).
30. Chong, J. J. H. *et al.* Progenitor cells identified by PDGFR- $\alpha$  expression in the developing and diseased human heart. *Stem Cells Dev.* **22**, 1932–1943 (2013).



**Extended Data Figure 1 | Cryopreservation does not affect hESC-CM engraftment.** **a**, Schematic representation of experimental design for cryopreservation testing experiments. **b**, Human genomes detected after

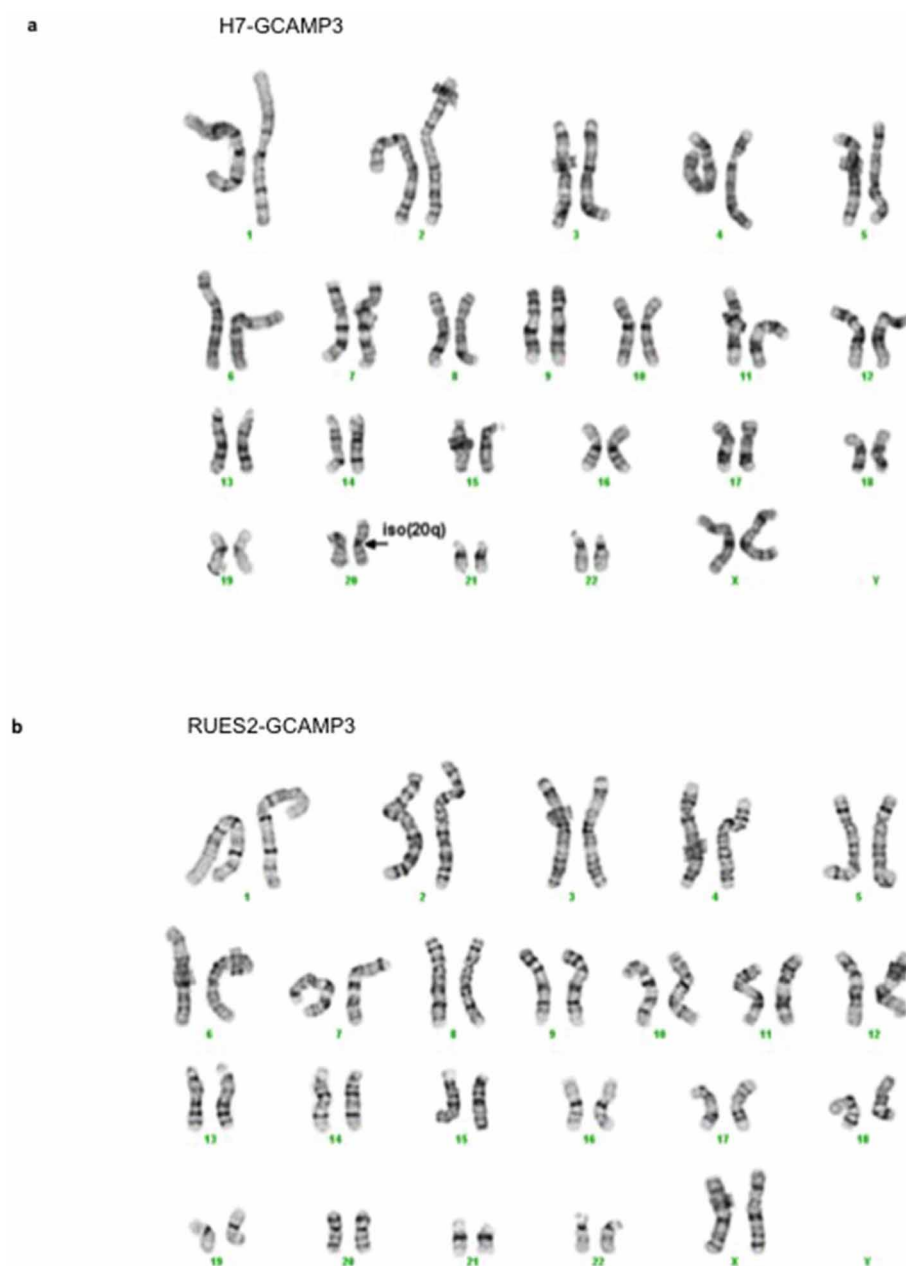
injection of cryopreserved or non-cryopreserved hESC-CMs were not significantly different ( $P > 0.05$ ,  $t$ -test). Mean  $\pm$  s.e.m. is shown ( $n = 9$  biological replicates) Experiment was performed once. NS, not significant.



**Extended Data Figure 2 | Creation and validation of the GCaMP3-expressing human ES-cell lines.** **a**, Targeting construct for ZFN engineering of GCaMP3 into the *AAVS1* locus. The endogenous genomic probe and neomycin resistance gene probe binding sites used for Southern blotting are shown.

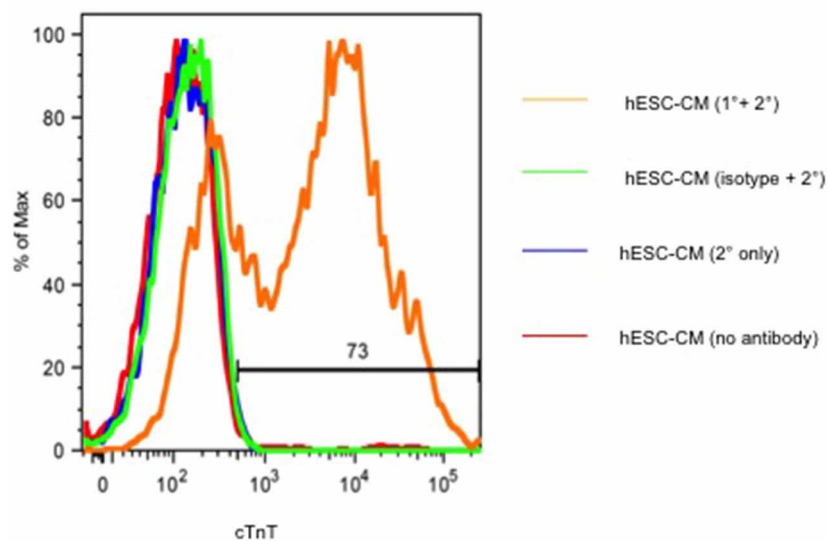
**b**, Southern blot analysis demonstrates a single integration event by hybridization for neomycin resistance cassette (left) and heterozygous *AAVS1* integration by genomic probe labelling (right).



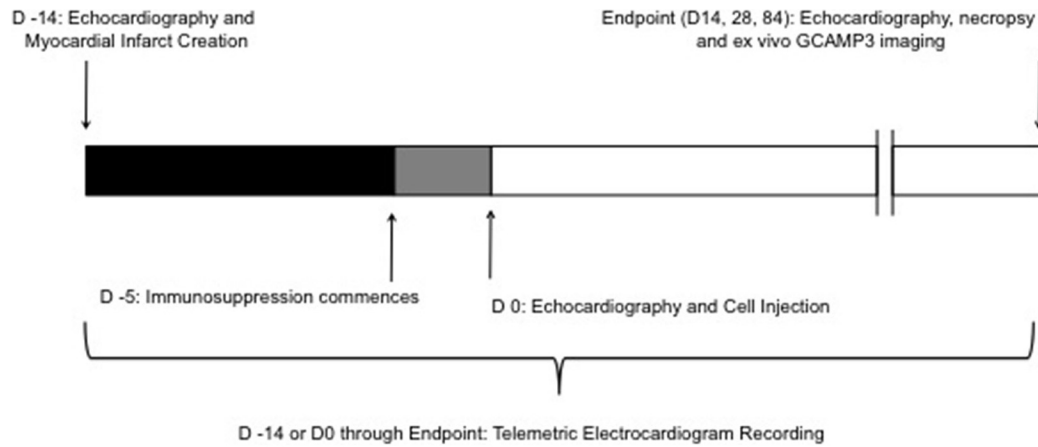


**Extended Data Figure 3 | Chromosomal analysis of human ES cells modified to encode GCaMP3.** a, H7-GCaMP3 ES-cell line demonstrates an

isochromosome of the chromosome 20 long arm (arrow). b, RUES2-GCaMP3 ES-cell line shows normal karyotype.



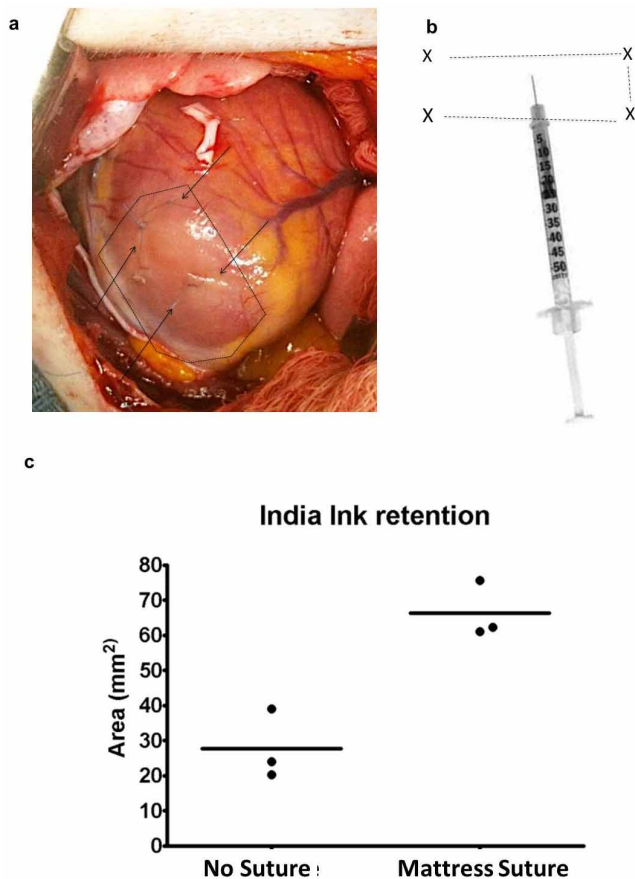
**Extended Data Figure 4 |** Flow cytometry for cardiomyocyte differentiation of human ES cells. Representative histogram of hESC-CMs after differentiation shows 73% cTnT-expressing cells.



**Extended Data Figure 5 | Schematic representation of experimental design.**

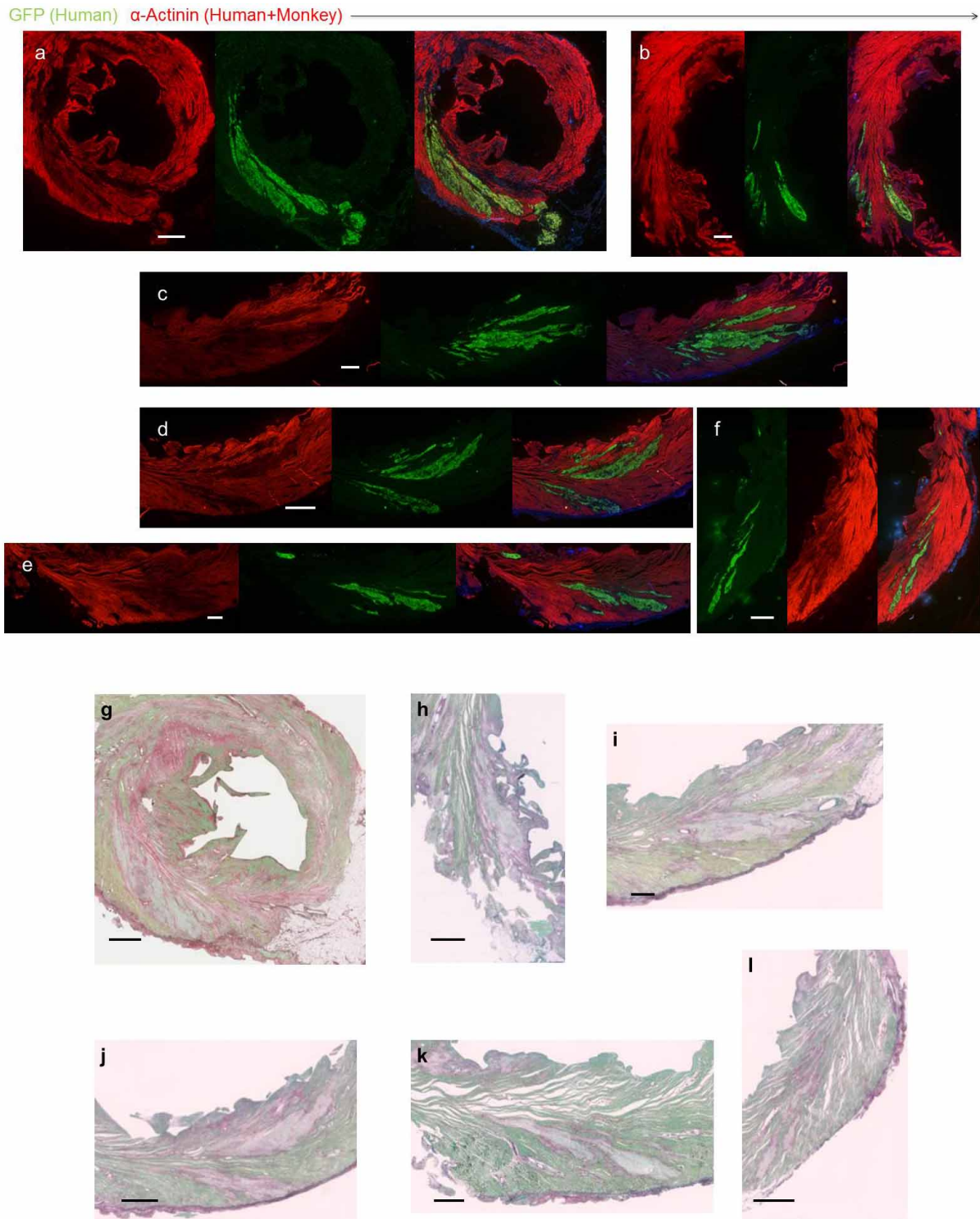
Myocardial infarction was created by advancing a balloon catheter into the distal left anterior descending artery and inflating it to create ischaemia (90 min) followed by reperfusion. The infarct was induced 14 days (D) before hESC-CM delivery via left thoracotomy. Immunosuppression using cyclosporine A, methylprednisolone and abatacept (T-cell co-stimulatory antagonizing fusion protein) was delivered 5 days before cell delivery and

continued until animals were euthanized. Primary endpoints were (1) histologically based morphometric calculations of infarct and graft size with analysis of graft composition, and (2) *ex vivo* analysis of graft–host electromechanical coupling enabled by GCaMP3 fluorescence detection. Secondary endpoints were (1) detection of arrhythmias by telemetric electrocardiogram analysis, and (2) analysis of left ventricular functional change by trans-oesophageal echocardiography.



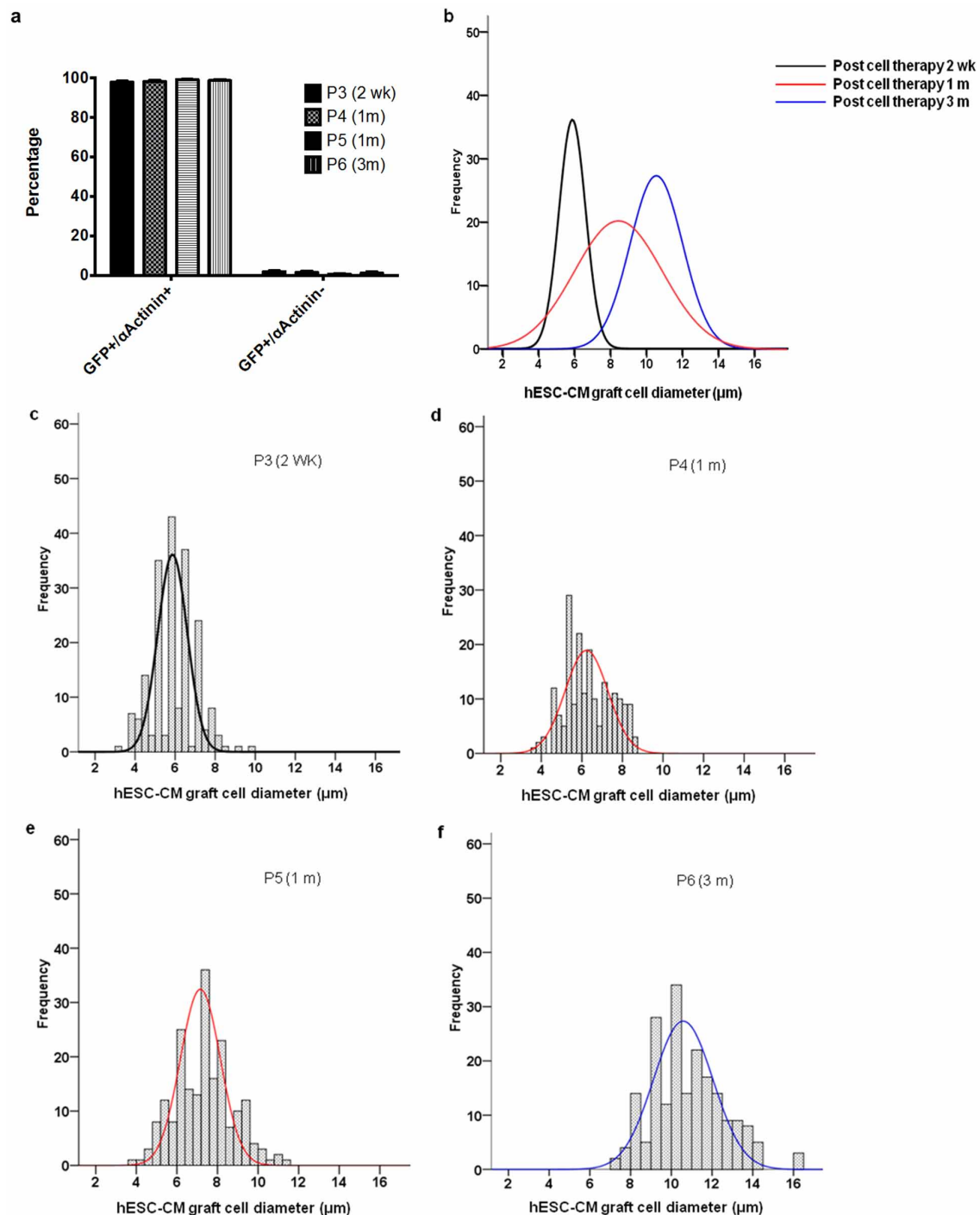
**Extended Data Figure 6 | Technique for hESC-CM injection to infarct region and border zones using 'mattress' suture strategy.** **a**, The macaque infarcted ventricular apex is seen as a blanched region (dotted line) during left thoracotomy. A total of 15 aliquots, each containing 100  $\mu$ l of hESC-CM in pro-survival cocktail, were delivered through five epicardial puncture sites (arrows, note one further puncture site not seen is on posterior aspect). **b**, hESC-CM retention after injection was increased by use of a 'mattress' suture. Crosses indicate insertion points of suture with dotted lines representing path of suture (exaggerated size for diagrammatic representation). A needle tip was inserted into the resulting rectangular area and the suture was tightened after a series of three injections (altering the trajectory of the needle) but before withdrawal of needle tip. **c**, Quantification of India ink retention after injection into left ventricular myocardium of anaesthetized macaques with or without use of the mattress suture technique ( $n = 3$  biological replicates each group). A trend favouring greater retention with the mattress suture is seen.





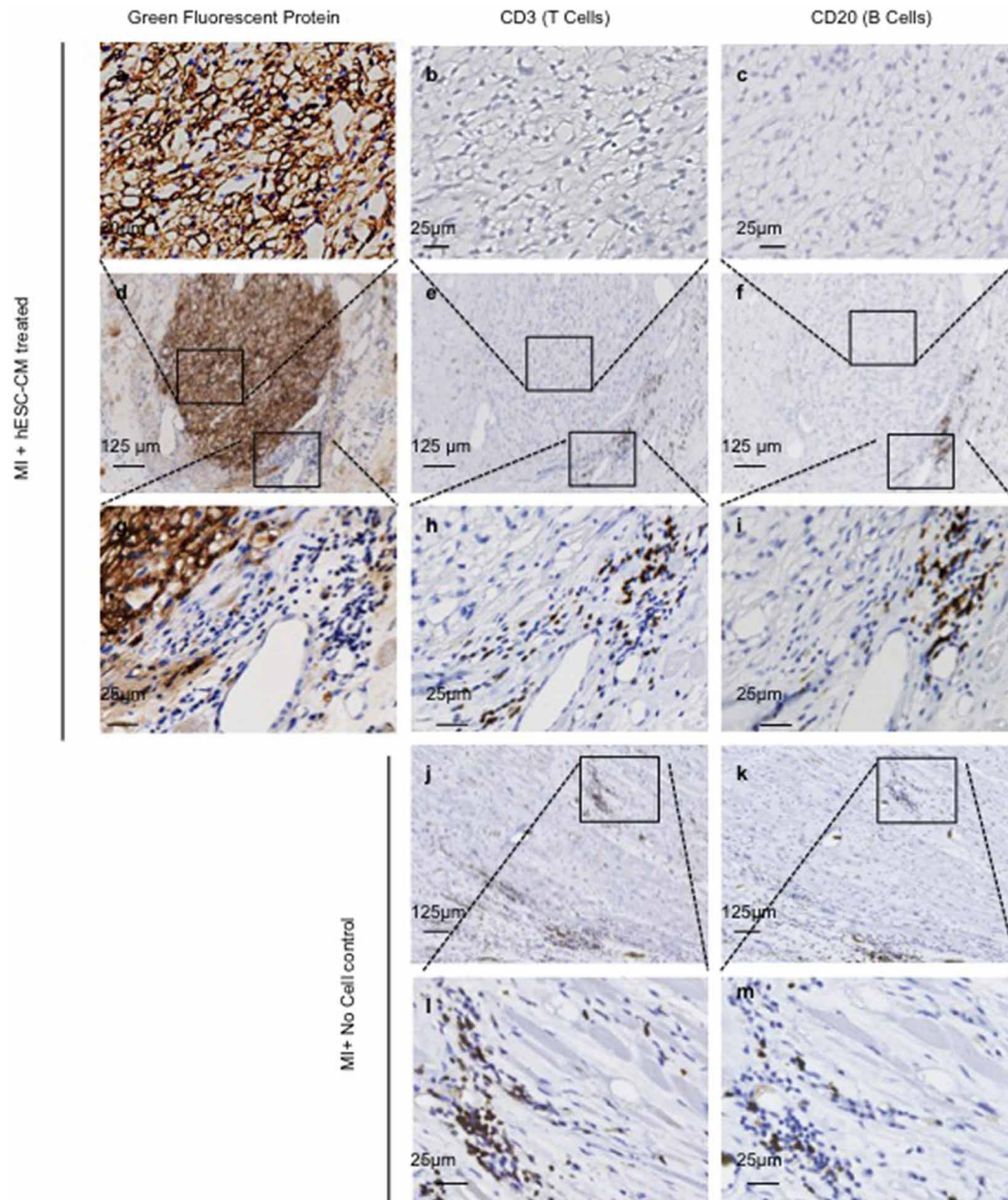
**Extended Data Figure 7 | Remuscularization of the infarcted macaque heart.** a–f, Single channels of confocal immunofluorescence shown in Fig. 1a–f. Macaque heart shown was subjected to myocardial infarction and transplantation of hESC-CMs 14 days before being euthanized. g–l, Picrosirius

Red staining of sections in close proximity to confocal immunofluorescence in a–f shows lack of fibrosis within hESC-CM grafts. Scale bars: 2,000  $\mu$ m (a, f, g, l); 1,000  $\mu$ m (b–e, h–k).



**Extended Data Figure 8 | Remuscularized infarct region is composed of engrafted cardiomyocytes that increase in size with time.** **a**, Quantification of the sarcomeric protein  $\alpha$ -actinin expression in GFP-expressing grafts. The vast majority (>98%) of GFP-expressing cells co-expressed  $\alpha$ -actinin. P3–P6 represent individual animals ( $n = 1$ ) euthanized at 2 weeks (2 wk) 1 month (m) or 3 months after hESC-CM delivery. Five hundred to seven hundred cells were counted from three different graft regions of each heart.

Percentage of GFP/ $\alpha$ -actinin double-positive cells and GFP-positive/ $\alpha$ -actinin-negative cells are shown as mean  $\pm$  s.d. **b**, Normal curve from histograms showing the distribution of human ES-cell-derived cardiomyocyte diameters (graft) in monkey hearts 2 weeks, 1 or 3 months after cell delivery. **c–f**, Individual histograms with superimposed normal curve of animals P3–P6 (as above).

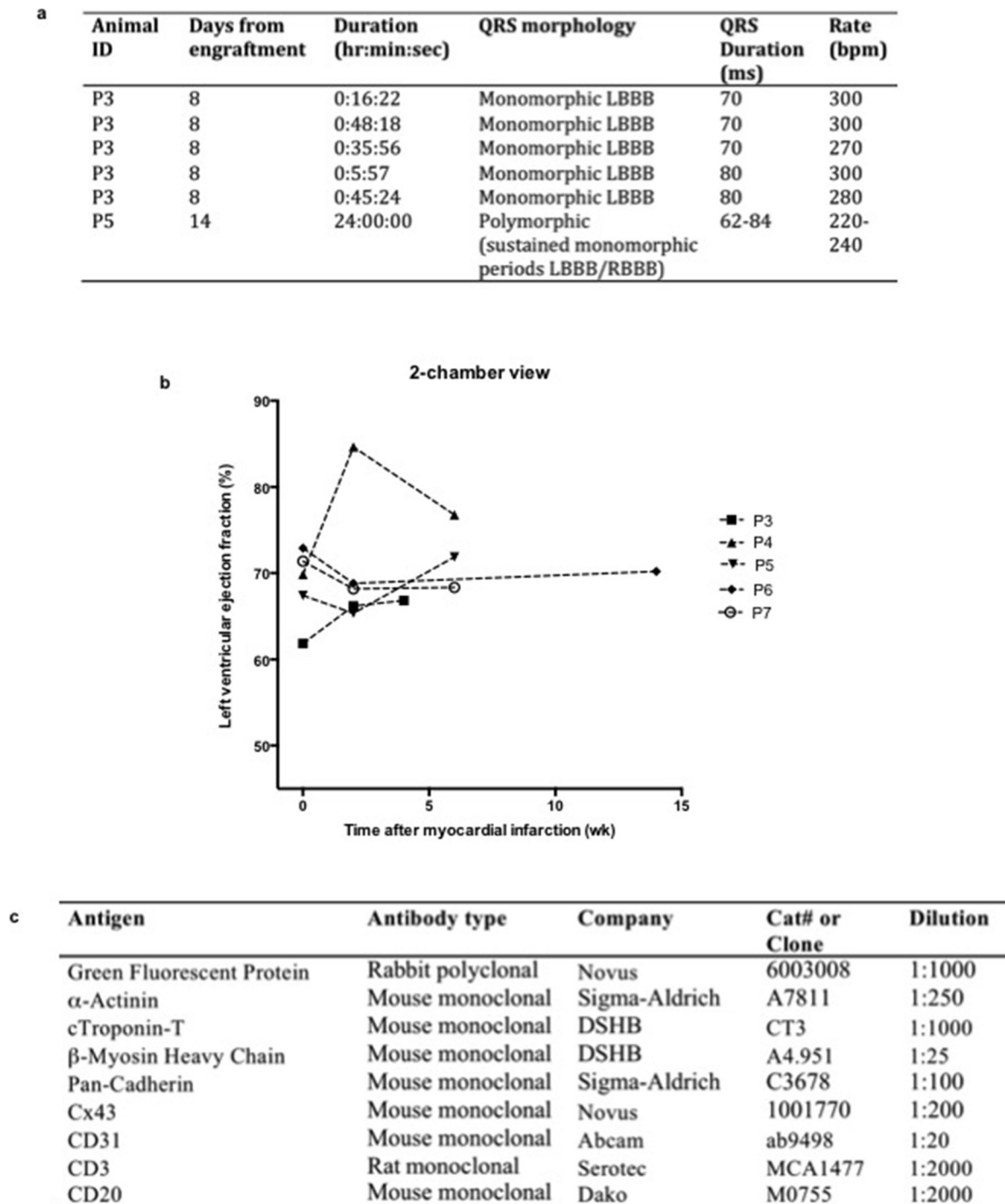


**Extended Data Figure 9 | No evidence of human graft rejection.**

**a–i**, Representative low- (**d–f**) and high- (**a–c** and **g–i**) power magnification of hESC-CM graft 28 days after cell delivery to infarcted macaque heart. Representative low- (**j–k**) and high- (**l, m**) power magnification of infarct region from control macaque 28 days after sham treatment. The hESC-CM

graft is detected by anti-GFP primary antibody with 3,3'-diaminobenzidine (DAB) detection of secondary antibody (brown). Few CD3<sup>+</sup> T lymphocytes or CD20<sup>+</sup> B lymphocytes are seen surrounding the hESC-CM grafts. Comparable numbers of T and B cells are seen in control infarcts receiving no hESC-CM treatment. Boxed inset regions show areas of higher magnification.





**Extended Data Figure 10 | Summary of ventricular tachycardia and echocardiographic assessment of left ventricular function.** **a**, Table characterizing episodes of ventricular tachycardia after engraftment of hESC-CMs (detailed in Fig. 5). Note that P5 demonstrated no discernible sinus rhythm on telemetric recording of the ECG 14 days after hESC-CM delivery. Although QRS morphology varied, the tachyarrhythmia comprised sustained periods of stable monomorphic QRS morphology. LBBB, left bundle branch block; RBBB, right bundle branch block. bpm, beats per minute. **b**, Left ventricular function was assessed by trans-oesophageal echocardiography at the following time points: before myocardial infarct creation, before hESC-CM delivery (2 weeks after myocardial infarction) and

before animals were euthanized (2, 4 or 12 weeks after myocardial infarction). P7 received no cells/vehicle only. All other animals received hESC-CMs. Results shown are for left ventricular ejection fraction calculated by two blinded cardiologists from the two-chamber view of the left ventricle. Note that this view best captures the infarcted antero-apical wall. The vehicle-treated control monkey showed a modest diminution in ejection fraction post-infarction. The cell-treated animals showed variable responses, with some having increased function and some having decreased function. Because of small group size, no statistical effects of hESC-CM therapy can be discerned. **c**, Table of antibodies used. DSHB, Developmental Studies Hybridoma Bank.



# Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer

Irfan A. Asangani<sup>1,2</sup>, Vijaya L. Dommeti<sup>1</sup>, Xiaoju Wang<sup>1,2</sup>, Rohit Malik<sup>1,2</sup>, Marcin Cieslik<sup>1</sup>, Rendong Yang<sup>3</sup>, June Escara-Wilke<sup>1</sup>, Kari Wilder-Romans<sup>4</sup>, Sudheer Dhanireddy<sup>1</sup>, Carl Engelke<sup>1</sup>, Mathew K. Iyer<sup>1</sup>, Xiaojun Jing<sup>1</sup>, Yi-Mi Wu<sup>1,2</sup>, Xuhong Cao<sup>1,5</sup>, Zhaohui S. Qin<sup>3</sup>, Shaomeng Wang<sup>6,7</sup>, Felix Y. Feng<sup>1,4,7</sup> & Arul M. Chinnaiyan<sup>1,2,5,7,8</sup>

**Men who develop metastatic castration-resistant prostate cancer (CRPC) invariably succumb to the disease. Progression to CRPC after androgen ablation therapy is predominantly driven by deregulated androgen receptor (AR) signalling<sup>1–3</sup>. Despite the success of recently approved therapies targeting AR signalling, such as abiraterone<sup>4–6</sup> and second-generation anti-androgens including MDV3100 (also known as enzalutamide)<sup>7,8</sup>, durable responses are limited, presumably owing to acquired resistance. Recently, JQ1 and I-BET762 two selective small-molecule inhibitors that target the amino-terminal bromodomains of BRD4, have been shown to exhibit anti-proliferative effects in a range of malignancies<sup>9–12</sup>. Here we show that AR-signalling-competent human CRPC cell lines are preferentially sensitive to bromodomain and extraterminal (BET) inhibition. BRD4 physically interacts with the N-terminal domain of AR and can be disrupted by JQ1 (refs 11, 13). Like the direct AR antagonist MDV3100, JQ1 disrupted AR recruitment to target gene loci. By contrast with MDV3100, JQ1 functions downstream of AR, and more potently abrogated BRD4 localization to AR target loci and AR-mediated gene transcription, including induction of the *TMPRSS2-ERG* gene fusion and its oncogenic activity. *In vivo*, BET bromodomain inhibition was more efficacious than direct AR antagonism in CRPC xenograft mouse models. Taken together, these studies provide a novel epigenetic approach for the concerted blockade of oncogenic drivers in advanced prostate cancer.**

The identification and therapeutic targeting of co-activators or mediators of AR transcriptional signalling should be considered as alternative strategies to treat CRPC<sup>14</sup>. BRD4 is a conserved member of the BET family of chromatin readers, which includes BRD2, BRD3 and BRDT. BRD4 has a critical role in transcription by RNA polymerase II (RNA Pol II) by facilitating recruitment of the positive transcription elongation factor P-TEFb<sup>15,16</sup>. Similar to other BET-family proteins, BRD4 contains two conserved bromodomains, BD1 and BD2. Competitive binding of JQ1 or I-BET762 to the bromodomain pocket results in the displacement of BRD4 from active chromatin and the subsequent removal of RNA Pol II from target genes<sup>10–13,17</sup>. Although most cancer cells express BET-family proteins, it is not clear why only a subset of cell lines from diverse cancers responds to BET inhibitors<sup>9,18</sup>. Recently, BRD4 was shown to interact with sequence-specific DNA-binding transcription factors in a gene-specific manner<sup>19</sup>. As the genetic and epigenetic landscape differs between tumour types, it is possible that distinct transcriptional regulators that associate with BRD4 might influence the action of BET inhibitors.

To discover new treatment options for CRPC, we treated a panel of five prostate cancer cell lines and one benign prostate cell line with JQ1, and found three of the AR-signalling positive cells to be sensitive to JQ1, although all six cell lines express high levels of its target proteins (Fig. 1a and Extended Data Fig. 1a, b). Next, knockdown of BRD2, 3 and 4 (Extended Data Fig. 1c) led to significant inhibition of cell proliferation and invasion, phenocopying JQ1 treatment (Extended Data Fig. 1d, e).

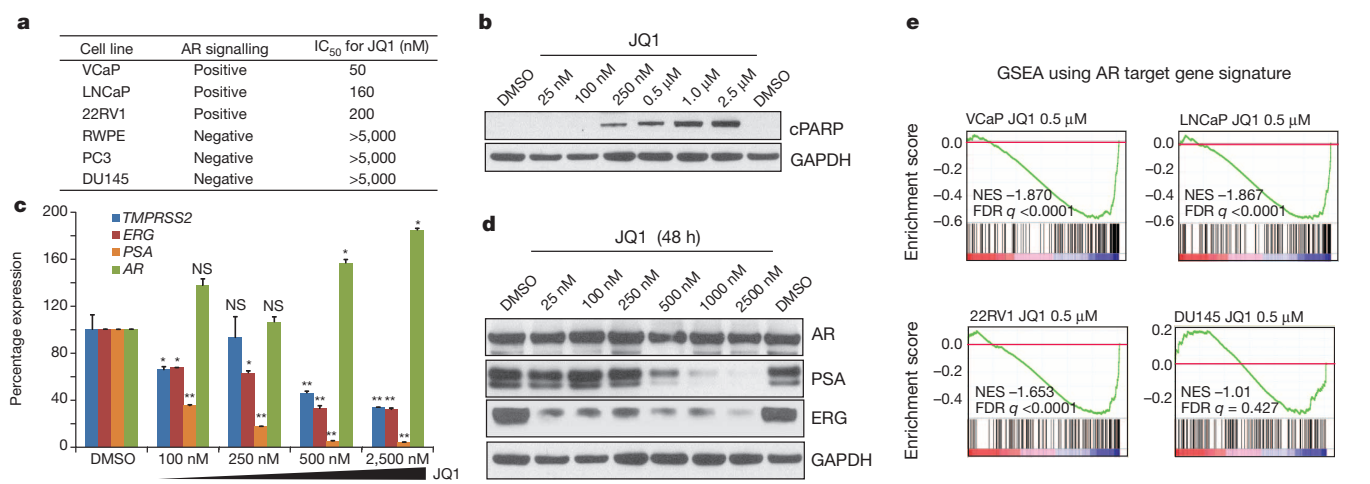
Furthermore, JQ1 treatment induced G<sub>0</sub>–G<sub>1</sub> arrest, apoptosis and associated transcriptional downregulation of the anti-apoptotic protein BCL-xl (also known as BCL2L1) in AR-positive cells<sup>13,18</sup> (Fig. 1b and Extended Data Fig. 1f–h). Similar to BCL2 downregulation by the BET inhibitor I-BET151 in leukaemia<sup>10</sup>, a reduction in BCL-xl by JQ1 could be explained in part by the observed loss of BRD2/3/4 recruitment to its promoter region (Extended Data Fig. 1j). Even at a relatively low 100 nanomolar (nM) concentration, long-term colony formation of AR-positive cells was severely inhibited by JQ1 (Extended Data Fig. 1k) with no apparent effect on JQ1 target proteins (Extended Data Fig. 1l, m).

As AR-positive cells were preferentially sensitive to JQ1, we examined whether JQ1 has an effect on AR target genes. VCaP human prostate cancer cells that harbour the *TMPRSS2-ERG* gene fusion and AR amplification<sup>20</sup> showed a dose-dependent decrease in prostate-specific antigen (PSA) and ERG at the messenger RNA and protein levels upon JQ1 treatment (Fig. 1d, e). Similar effects were observed in LNCaP and 22RV1 prostate cancer cells (Extended Data Fig. 2a, b). Furthermore, bortezomib did not reverse the JQ1-mediated PSA and ERG protein loss, indicating that these genes are regulated at the transcriptional level (Extended Data Fig. 2c).

We performed microarray analysis to examine changes in global gene expression upon JQ1 treatment. Gene set enrichment analysis (GSEA) using the AR gene signature revealed significant repression of these genes in AR-positive cells (Fig. 1f), suggesting a role of BET proteins in AR-mediated transcription. Additionally, we observed a loss of the MYC-associated gene signature in AR-positive cell lines upon JQ1-treatment (Extended Data Fig. 2d). MYC is a known transcriptional target of BET inhibition in haematological cancers<sup>11,18</sup>. Interestingly, MYC levels were attenuated by JQ1 in cells that are AR positive and sensitive to JQ1 inhibition, but not in AR-negative cells (Extended Data Fig. 2e). Thus, high expression of MYC *per se* (Extended Data Fig. 1b) does not confer sensitivity to JQ1 in prostate cancer cells. Time-course experiments with JQ1 demonstrated loss of MYC (Extended Data Fig. 2f, g) and cyclohexamide had no additional effect on MYC protein levels (Extended Data Fig. 2h, i), ruling out a post-translational mode of JQ1 action. Phenotypically, knockdown of MYC did not affect cell invasion (Extended Data Fig. 2j), whereas JQ1 treatment inhibited invasion (Extended Data Fig. 1e). Furthermore, exogenous expression of MYC did not result in rescue of JQ1-mediated inhibition of cell growth (Extended Data Fig. 2k, l). Thus, although MYC levels may be repressed by JQ1 in AR-positive cells, and may have a role in proliferation, MYC does not seem to be the primary target for the antineoplastic effects of JQ1.

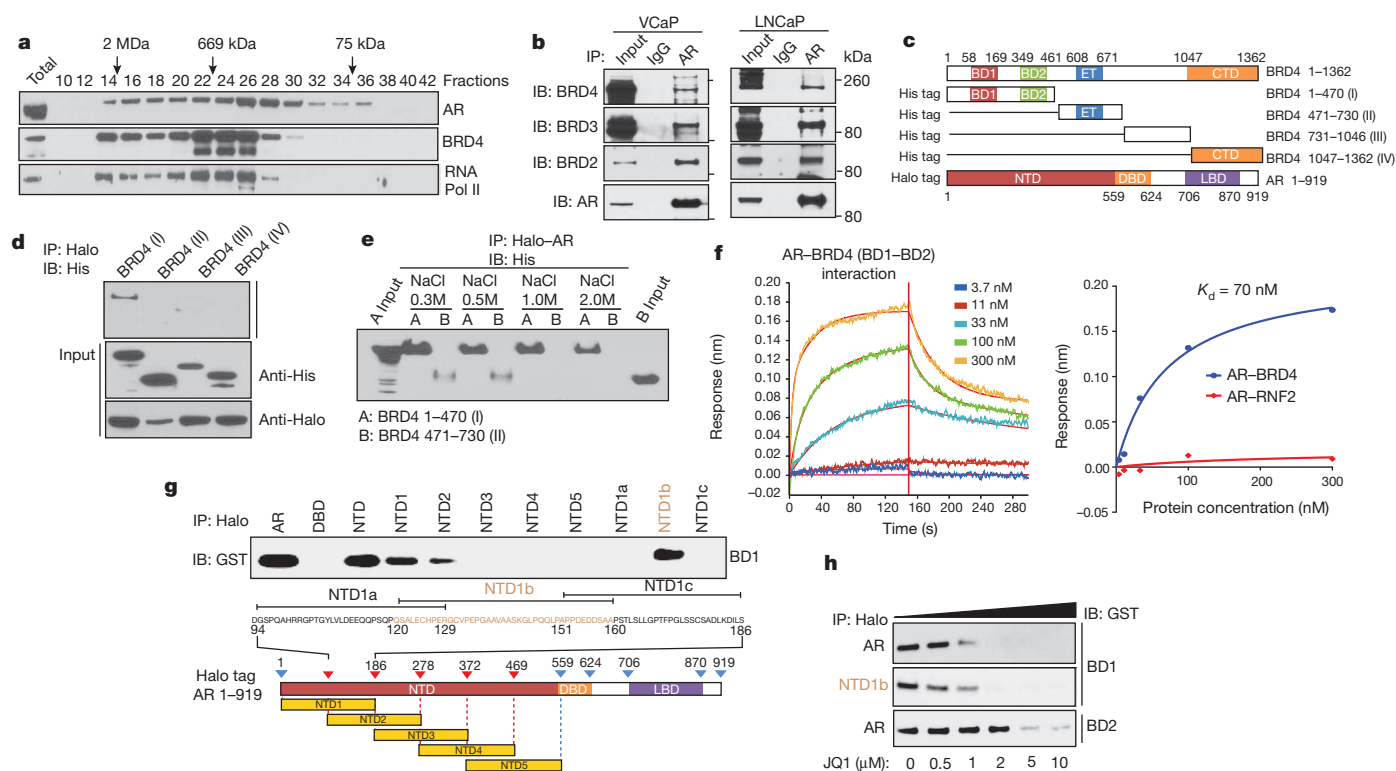
As BRD4 is known to engage sequence-specific DNA-binding proteins<sup>19</sup>, we proposed that AR may interact directly with BRD4. We performed gel-filtration chromatography and found that AR and BRD4 predominantly co-eluted in a high-molecular-weight complex (Fig. 2a and Extended Data Fig. 3a). Moreover, RNA Pol II, a reported target for phosphorylation by

<sup>1</sup>Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>2</sup>Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>3</sup>Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia 30329, USA. <sup>4</sup>Department of Radiation Oncology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>5</sup>Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>6</sup>Departments of Internal Medicine, Pharmacology, and Medicinal Chemistry, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>7</sup>Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. <sup>8</sup>Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.



**Figure 1 | Prostate cancer cell lines with intact androgen signalling are sensitive to BET bromodomain inhibition.** **a**, Half-maximum inhibitory concentration (IC<sub>50</sub>) for JQ1 in each cell line is shown. **b**, Induction of apoptosis in VCaP prostate cancer cells by JQ1. Cleaved PARP (cPARP) immunoblot analysis. GAPDH served as a loading control. DMSO, dimethylsulphoxide. **c**, Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

analysis of indicated genes in VCaP cells treated with varying concentrations of JQ1 for 24 h. Data represent mean  $\pm$  standard error of the mean (s.e.m.) ( $n = 3$ ) from one of three independent experiments. **d**, Immunoblot analyses of AR, PSA and ERG levels in VCaP cells treated with JQ1. **e**, GSEA of the AR target gene signature in VCaP, LNCaP, 22RV1 and DU145 cells. NS, not significant. \* $P \leq 0.05$ , \*\* $P \leq 0.005$  by two-tailed Student's  $t$ -test.



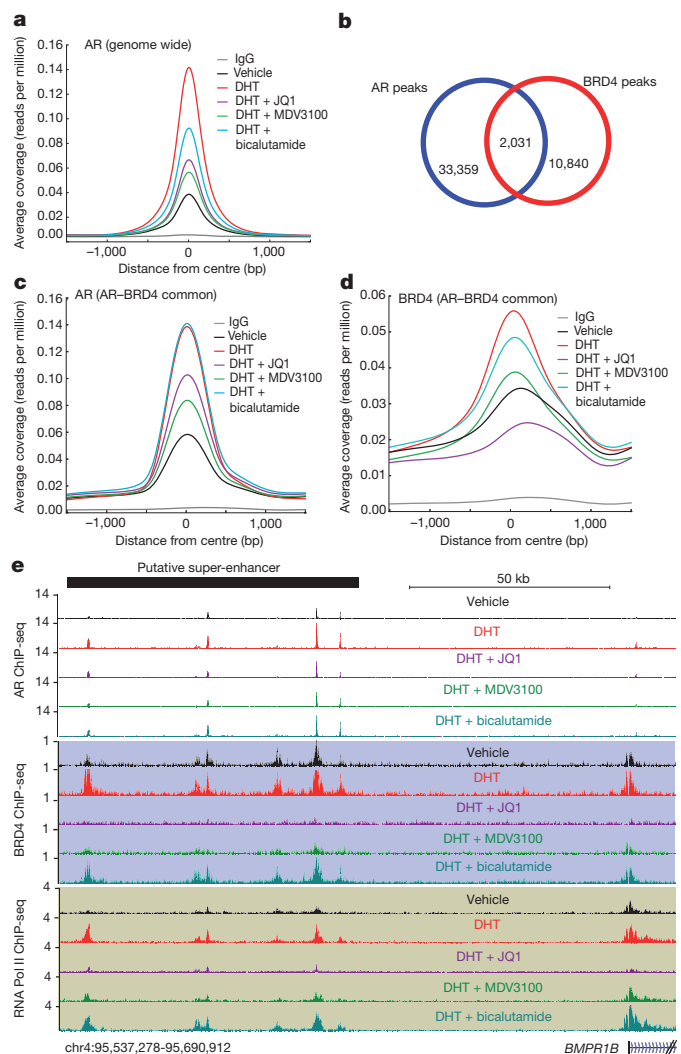
**Figure 2 | Physical association of the N-terminal domain of AR with BRD4 and its disruption by BET bromodomain inhibition.** **a**, VCaP nuclear extracts were fractionated on a Superose-6 column and AR, BRD4 and RNA Pol II were analysed by immunoblotting. **b**, Endogenous association of AR and BRD2/3/4. VCaP and LNCaP nuclear extracts were subjected to immunoprecipitation using an anti-AR antibody. Immunoprecipitates (IPs) were analysed for the presence of BRD2/3/4 by immunoblotting (IB; top). The immunoblot was stripped and reprobed for AR (bottom). 5% total lysate was used as input control. **c**, Schematic of BRD4 and AR constructs used for co-immunoprecipitation experiments. CTD, C-terminal domain; ET, extraterminal domain; DBD, DNA-binding domain; LBD, ligand-binding domain. **d**, NTD of BRD4 interacts with AR. Proteins from 293T cells co-transfected with various His-BRD4 deletions and Halo-AR constructs were subjected to immunoprecipitation with Halo beads followed by immunoblotting with His-tag antibody. Inputs are shown in the bottom panel.

**e**, As in **d** but with the indicated salt concentrations. **f**, Representative sensorgrams from three independent experiments for AR-BRD4 (BD1-BD2) by an OctetRED biolayer interferometry showing direct interaction. Real-time binding was measured by immobilizing biotinylated AR protein on the super streptavidin biosensor and subsequent interaction with varying concentrations of BRD4 (BD1-BD2) protein. The plots show the response versus protein concentration curves derived from the raw binding data. Right,  $K_d$  represents the BRD4 (BD1-BD2) concentration yielding half-maximal binding to AR. Protein RNF2 was used as negative control. **g**, NTD domain of AR interacts with BD1 of BRD4. Equal amounts of *in vitro* translated proteins were combined and immunoprecipitated using Halo beads followed by immunoblot analysis with anti-GST antibody. **h**, JQ1 disrupts AR-BD1 interactions. Varying concentrations of JQ1 were incubated with AR-BD1, NTD1b-BD1 and AR-BD2 complexes before immunoprecipitation followed by immunoblot analysis.

BRD4 (ref. 21), also co-eluted in the same complex, suggestive of a large multi-protein complex composed of AR, BRD4 and RNA Pol II. Immunoprecipitation experiments further confirmed an endogenous association between AR and BRD4 (Fig. 2b). Additionally, we observed an interaction between AR and BRD2/3 (Fig. 2b), implicating a common region in BRD2, 3 and 4 proteins responsible for AR interaction. To map the region mediating this interaction, we tested the ability of different deletion variants of BRD4 to pull-down AR in 293T cells (Fig. 2c). A BRD4 variant containing the BD1 and BD2 domains maintained the ability to pull-down AR even at high salt concentrations (Fig. 2d, e). To determine whether the BD1–BD2 domains directly interact with AR, we carried out quantitative assessment of the binding affinity using the OctetRED system. We applied varying concentrations of BD1–BD2 protein to biosensors with immobilized AR and found that BRD4 interacts with AR in a concentration-dependent fashion, with an estimated dissociation constant ( $K_d$ ) of 70 nM, supporting a high-affinity interaction (Fig. 2f and Extended Data Fig. 3b, c). To fine map this interaction we created a series of Halo-tagged AR and glutathione S-transferase (GST)-tagged BRD4 constructs for *in vitro* pull-down studies and demonstrated that the BD1, and to a lesser extent the BD2 domain, bind directly to the N-terminal domain (NTD) domain of AR, which was further mapped to a 38-amino-acid region—NTD1b—of AR (Fig. 2g and Extended Data Fig. 3d–f). Subsequently, we observed the disruption of BD1–AR and BD1–NTD1b interactions by JQ1 (Fig. 2h), as well as loss of endogenous BRD4–AR interactions (Extended Data Fig. 3g). Together, these data indicate that BET protein inhibition leads to disruption of the AR–BRD4 interaction, which probably explains the preferential activity of JQ1 in AR-positive prostate cancer cells.

The ubiquitously expressed BRD2, 3 and 4 proteins are suggested to have overlapping functions<sup>10,13,22</sup> and, consistent with this notion, we observed AR interactions with all three. Because BET inhibitors, such as JQ1 and I-BET762, have high affinity for the BD1 and BD2 domains of BRD2/3/4 proteins<sup>10,11,13</sup>, we proposed that BET inhibitors may affect genome-wide recruitment of all three BET proteins. We performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) with antibodies against BRD2/3/4 in VCaP cells treated with JQ1 or I-BET762 (Extended Data Fig. 4a) and observed a high genome-wide overlap between BRD2, 3 and 4 (62–86% peak overlap) (Extended Data Fig. 4b, c). JQ1 or I-BET762 treatment led to a reduction in the recruitment of all three proteins to chromatin (Extended Data Fig. 4d). Moreover, the reduced BRD2/3/4 recruitment was equally distributed for regions with or without AR (Extended Data Fig. 4e).

Binding of androgen (dihydrotestosterone (DHT)) to AR leads to its translocation from the cytoplasm to the nucleus, where it binds to regions of DNA harbouring androgen-responsive elements (AREs) and results in subsequent recruitment of proteins involved in transcriptional activation or suppression in a gene-specific manner. BRD4 interacts with acetylated histones as well as DNA-binding transcription factors, leading to context-dependent transcriptional activation or inhibition of target genes<sup>15,19,22</sup>. As the AR–BRD4 interaction is disrupted by JQ1 (Fig. 2), we next explored whether AR localization is affected in a genome-wide context. We performed ChIP-seq with antibodies against AR, BRD4 and RNA Pol II in cells that were either starved, treated with DHT or DHT plus JQ1 (Extended Data Fig. 4a). Two anti-androgens, bicalutamide and MDV3100, were included for comparison. As expected, the average ChIP-seq signal for AR was highly enriched in DHT-treated cells (Fig. 3a and Extended Data Fig. 5a, b). Recruitment of AR to target loci was markedly attenuated by MDV3100 and less so by bicalutamide. Interestingly, JQ1 blocked AR recruitment almost as effectively as MDV3100 (Fig. 3a and Extended Data Fig. 5c–e). Furthermore, we observed a co-recruitment of AR and BRD4 at 2,031 sites. The strongest association was observed within promoters of AR-regulated genes (502 promoters,  $P = 4 \times 10^{-49}$ ), and for the highest AR peaks (1,112 sites,  $P = 1 \times 10^{-38}$ ) (Fig. 3b). Limiting our evaluation to AR and BRD4 coincident peaks, we observed that DHT-mediated AR recruitment to these loci was inhibited by MDV3100 and to a lesser extent by JQ1 (Fig. 3c). By contrast,



**Figure 3 | BET bromodomain inhibition disrupts AR and BRD4 binding to target loci.** **a**, AR ChIP-seq was performed in VCaP cells treated for 12 h with vehicle, DHT (10 nM), DHT plus JQ1 (500 nM), DHT plus MDV3100 (10  $\mu$ M) or DHT plus bicalutamide (25  $\mu$ M). Summary plot of AR enrichment (average coverage) across AR-binding sites (ARs) in different treatment groups is shown. Data represent one of two biological replicates. **b**, Venn diagram illustrating the overlap of AR- and BRD4-enriched peaks in DHT-treated sample. **c**, **d**, Summary plot for AR and BRD4 enrichment for the AR–BRD4 overlapping (2,031) regions. **e**, Genome browser representation of AR, BRD4 and RNA Pol II binding events on a putative super-enhancer of the AR-regulated *BMPRI1B* gene. The y-axis denotes reads per million per base pair. The x-axis denotes the genomic position with a scale bar on top right. The putative super-enhancer region enriched for AR, BRD4 and RNA Pol II is depicted with a black bar on the top left. chr4, chromosome 4.

JQ1 almost completely abrogated DHT-induced BRD4 recruitment to the AR–BRD4 shared loci (Fig. 3d). Examples of gene tracks for AR- and BRD4-associated genomic regions such as enhancers and super-enhancers<sup>17</sup> and the effects of different treatments on their levels are shown in Fig. 3e and Extended Data Fig. 5f. Corroborating the ChIP-seq data, gene expression analysis in VCaP and LNCaP cells showed more efficient repression of DHT-induced AR-target genes by JQ1 than by MDV3100 or bicalutamide (Extended Data Fig. 5g, h).

JQ1 treatment had a marked effect on ERG expression in VCaP cells (Fig. 1d, e and Extended Data Fig. 5h), and we found that the attenuation of DHT-induced ERG expression by JQ1 was due to de-recruitment of RNA Pol II from the *ERG* gene body and reduced binding of AR and BRD4 at the *TPRSS2* promoter/enhancer (Extended Data Fig. 6a, b). The efficient downregulation of ERG by JQ1 has important implications,



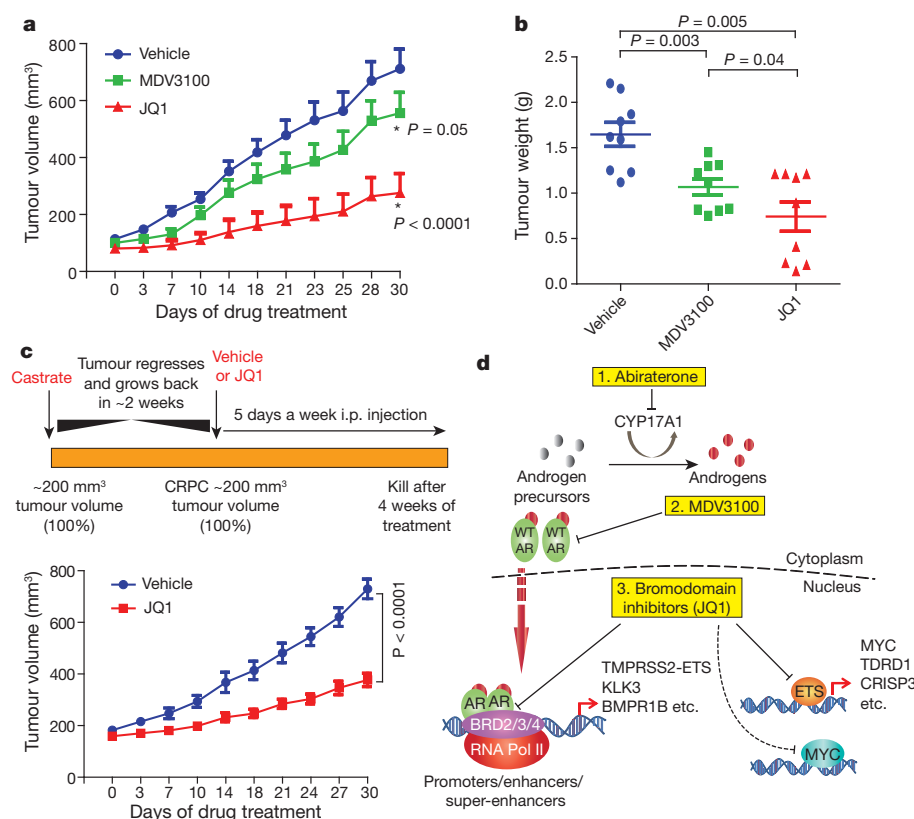
as the *TMPRSS2-ERG* gene fusion product is the oncogenic driver in 50% of prostate cancers<sup>20,23</sup>. To investigate the effect of JQ1 on ERG-mediated transcription, we performed ERG ChIP-seq in cells treated with JQ1 for 12 h—a time window in which ERG protein levels are still unaffected by JQ1 (Extended Data Fig. 6c)—and observed a significant loss in the top 4% of ERG enriched peaks (Extended Data Fig. 6d).

We next determined the functional consequences of JQ1 treatment by measuring the expression levels of select ERG target genes (Extended Data Fig. 6e, f). As expected, the ERG-activated genes were downregulated and the ERG-repressed genes were de-repressed by JQ1 (Extended Data Fig. 6g, h). To evaluate BET inhibitor repression of ERG-mediated oncogenic function in an isogenic setting, we overexpressed ERG in RWPE and PC3 cells (Extended Data Fig. 7a, b). Treatment with JQ1 or I-BET762 led to an attenuation of ERG-mediated invasion (Extended Data Fig. 7c) and GSEA demonstrated a significant negative enrichment for ERG target genes upon BET inhibitor treatment (Extended Data Fig. 7d). Furthermore, we found that ERG was highly enriched on the known distal enhancer of *MYC* that was reduced upon JQ treatment (Extended Data Fig. 8a, b). Likewise, ETV1 occupies the same distal enhancer region in ETV1 fusion-positive LNCaP cells<sup>23</sup>. Knockdown of ERG or ETV1 along with AR led to *MYC* downregulation, implicating *MYC* regulation by ETS proteins in fusion-positive prostate cancer cells (Extended Data Fig. 8c–e). Notably, ChIP-seq analysis of AR and RNA Pol II enrichment at the *MYC* locus presented an interesting pattern in which DHT treatment led to increased AR and reduced RNA Pol II binding on the *MYC* distal enhancer and gene body, respectively, which was reinstated in the presence of MDV3100 or bicalutamide but not JQ1 (Extended Data Fig. 8f). This observation is consistent with the concomitant reduction in *MYC* expression upon DHT treatment that was de-repressed in the presence of MDV3100 but not JQ1 (Extended Data Fig. 8g–i). Lack of de-repression of *MYC* by JQ1 in this setting could be explained by the fact that both AR and ERG are absent from the *MYC* distal enhancer, leading to net loss of *MYC* expression. These data also suggest a mechanism by which CRPC patients become resistant to anti-androgen therapy by maintaining expression of the *MYC* oncogene.

Next, we sought to compare the efficacy of JQ1 and MDV3100, a direct AR antagonist used clinically to treat advanced CRPC<sup>8</sup>. Before embarking on the *in vivo* experiment we tested them on VCaP cells *in vitro* for 8 days and observed marginal cell death by MDV3100 versus suppression of cell growth at sub-micromolar concentrations of JQ1 (Extended Data Fig. 9a). To rule out the possibility of JQ1 being a generic anti-androgen we confirmed that JQ1 had no effect on physiological androgen-regulated processes; however, JQ1 reduced testes size in mice, as reported previously<sup>24</sup> (Extended Data Fig. 9b–f). Treatment of VCaP tumour-bearing mice with JQ1 led to a significant reduction in tumour volume/weight (Fig. 4a, b and Extended Data Fig. 10a), whereas MDV3100 had a less pronounced effect.

Recently, several studies described the pro-metastatic effects of MDV3100 in pre-clinical models<sup>25</sup>. To test whether MDV3100 treatment leads to spontaneous metastasis in our VCaP xenograft model, we isolated femur, liver and spleen from MDV3100-treated mice and found evidence of metastases in femur and liver (Extended Data Fig. 10b, c). By contrast, JQ1-treated mice showed no evidence of metastasis (Extended Data Fig. 10c). Taken together, these pre-clinical studies suggest that the use of MDV3100 in clinically localized prostate cancer may potentiate the formation of micro-metastases, unlike BET inhibitors. Consistent with previous reports, JQ1 and MDV3100 were both well tolerated by mice (Extended Data Fig. 10d). Although VCaP cells were originally derived from a patient with CRPC, VCaP tumour xenografts respond to castration in mouse models. We found that JQ1 still had a growth inhibitory effect in castration-resistant VCaP tumour xenografts and observed a 50% reduction in castration-resistant tumours by JQ1-treatment (Fig. 4c and Extended Data Fig. 10e).

Maintenance of AR signalling is the most common resistance mechanism that patients with advanced prostate cancer develop after conventional hormonal treatments<sup>26</sup>. AR amplification, mutation and alternative splicing have all been suggested as potential resistance mechanisms to anti-androgen treatments<sup>2,27,28</sup>. Over half of CRPC patients have at least one of these aberrations in the AR pathway<sup>29</sup>. As BET inhibitors function ‘downstream’ of AR (Fig. 4d), our data indicate that these compounds may be effective in the context of AR-mediated resistance, including



**Figure 4 | BET bromodomain inhibition blocks CRPC *in vivo*.** **a**, VCaP cells were implanted subcutaneously in mice and grown until tumours reached a size of approximately 100 mm<sup>3</sup>. Xenografted mice were randomized and then received ( $n = 6$  per group) vehicle, 50 mg kg<sup>-1</sup> JQ1 or 10 mg kg<sup>-1</sup> MDV3100, as indicated, 5 days a week. Calliper measurements were taken twice a week. Mean tumour volume  $\pm$  s.e.m. is shown. **b**, Individual tumour weight from different treatment groups with  $P$  values is shown. **c**, Top, schematic illustrating the VCaP CRPC mouse xenograft experimental design. i.p., intraperitoneal. Bottom, castrated mice bearing the VCaP CRPC xenograft received vehicle ( $n = 6$ ) or 50 mg kg<sup>-1</sup> JQ1 ( $n = 7$ ), as indicated, 5 days a week. Mean tumour volume  $\pm$  s.e.m. is shown. Statistical significance was determined by two-tailed Student's *t*-test. **d**, Schematic depicting varying mechanisms to block AR signalling in CRPC. First, abiraterone inhibits androgen biosynthesis by blocking the enzyme CYP17A1. Second, MDV3100 competitively antagonizes androgen binding to AR preventing nuclear translocation and recruitment to target gene loci. Third, JQ1 (or BET inhibitors) blocks AR and BRD2/3/4 interaction and co-recruitment to target gene loci as well as the functional activity and/or expression of ETS and MYC. WT, wild type.



compensatory mechanisms involving related steroid hormone receptors that are also likely to require BET bromodomain function. By functioning downstream of AR, BET inhibition is less likely to be affected by acquired resistance associated with AR antagonists, including the recently identified F876L mutation of AR<sup>30</sup>. Although both MDV3100 and JQ1 block AR recruitment to target loci on a genome-wide scale (the 'AR cistrome'), we found that JQ1 probably has an enhanced inhibitory effect by further abrogating co-recruitment of BRD4, which is required for mobilization of the transcriptional machinery<sup>15,16</sup>.

A recent study demonstrated that BET inhibition leads to preferential loss of BRD4 at super-enhancers and consequent transcriptional elongation defects<sup>17</sup>. These super-enhancers were often associated with key oncogenic drivers in a variety of cancers. Tumour cells are thought to become addicted to select oncogenes and hence unusually reliant on their high expression, which may explain the preferential sensitivity of BET inhibition in cancer versus normal tissues. Although MYC and its association with multiple myeloma was highlighted as a super-enhancer-dependent cancer<sup>17</sup>, this framework probably applies to key transcription factors involved in the development of CRPC, including AR, ETS and MYC (Fig. 4d). Taken together, these data indicate that clinical evaluation of BET inhibitors is warranted in CRPC, either as monotherapy or in combination with second-generation anti-androgens.

## METHODS SUMMARY

Gene expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray following the manufacturer's protocol. Proteins were extracted by lysing the cells in RIPA lysis buffer (Sigma) supplemented with protease inhibitor cocktail (Sigma). Immunoblotting was performed with standard protocols using polyvinylidene difluoride (PVDF) membrane (GE Healthcare), and signals were visualized with an enhanced chemiluminescence system as described by the manufacturer (GE Healthcare). The ChIP assays for BRD2, BRD3, BRD4, AR, RNA Pol II, ERG and H3K27ac with specific antibodies were performed using HighCell ChIP kit (Diagenode) following the manufacturer's protocol. All procedures involving mice were approved by the University Committee on Use and Care of Animals at the University of Michigan and conform to all regulatory standards.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 May 2013; accepted 5 March 2014.

Published online 23 April 2014.

1. Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).
2. Chen, C. D. *et al.* Molecular determinants of resistance to antiandrogen therapy. *Nature Med.* **10**, 33–39 (2004).
3. Visakorpi, T. *et al.* *In vivo* amplification of the androgen receptor gene and progression of human prostate cancer. *Nature Genet.* **9**, 401–406 (1995).
4. Stein, M. N., Goodin, S. & Dapaola, R. S. Abiraterone in prostate cancer: a new angle to an old problem. *Clin. Cancer Res.* **18**, 1848–1854 (2012).
5. Reid, A. H. *et al.* Significant and sustained antitumor activity in post-docetaxel, castration-resistant prostate cancer with the CYP17 inhibitor abiraterone acetate. *J. Clin. Oncol.* **28**, 1489–1495 (2010).
6. de Bono, J. S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* **364**, 1995–2005 (2011).
7. Mukherji, D., Pezaro, C. J. & De Bono, J. S. MDV3100 for the treatment of prostate cancer. *Expert Opin. Investig. Drugs* **21**, 227–233 (2012).
8. Scher, H. I. *et al.* Increased survival with enzalutamide in prostate cancer after chemotherapy. *N. Engl. J. Med.* **367**, 1187–1197 (2012).
9. Lockwood, W. W., Zejnullahu, K., Bradner, J. E. & Varmus, H. Sensitivity of human lung adenocarcinoma cell lines to targeted inhibition of BET epigenetic signaling proteins. *Proc. Natl Acad. Sci. USA* **109**, 19408–19413 (2012).
10. Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).

11. Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904–917 (2011).
12. Puissant, A. *et al.* Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
13. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
14. Attard, G., Richards, J. & de Bono, J. S. New strategies in metastatic prostate cancer: targeting the androgen receptor signaling pathway. *Clin. Cancer Res.* **17**, 1649–1657 (2011).
15. Jang, M. K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
16. Yang, Z. *et al.* Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535–545 (2005).
17. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
18. Mertz, J. A. *et al.* Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl Acad. Sci. USA* **108**, 16669–16674 (2011).
19. Wu, S. Y., Lee, A. Y., Lai, H. T., Zhang, H. & Chiang, C. M. Phospho-switch triggers Brd4 chromatin binding and activator recruitment for gene-specific targeting. *Mol. Cell* **49**, 843–857 (2013).
20. Tomlins, S. A. *et al.* Recurrent fusion of *TMPRSS2* and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648 (2005).
21. Devaiah, B. N. *et al.* BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. *Proc. Natl Acad. Sci. USA* **109**, 6927–6932 (2012).
22. Belkina, A. C. & Denis, G. V. BET domain co-regulators in obesity, inflammation and cancer. *Nature Rev. Cancer* **12**, 465–477 (2012).
23. Chen, Y. *et al.* ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. *Nature Med.* **19**, 1023–1029 (2013).
24. Matzuk, M. M. *et al.* Small-molecule inhibition of BRDT for male contraception. *Cell* **150**, 673–684 (2012).
25. Lin, T. H. *et al.* Anti-androgen receptor ASC-J9 versus anti-androgens MDV3100 (Enzalutamide) or Casodex (Bicalutamide) leads to opposite effects on prostate cancer metastasis via differential modulation of macrophage infiltration and STAT3–CCL2 signaling. *Cell Death Dis.* **4**, e764 (2013).
26. Harris, W. P., Mostaghel, E. A., Nelson, P. S. & Montgomery, B. Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nat. Clin. Pract. Urol.* **6**, 76–85 (2009).
27. Taplin, M. E. *et al.* Selection for androgen receptor mutations in prostate cancers treated with androgen antagonist. *Cancer Res.* **59**, 2511–2515 (1999).
28. Sun, S. *et al.* Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *J. Clin. Invest.* **120**, 2715–2730 (2010).
29. Grasso, C. S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).
30. Balbas, M. D. *et al.* Overcoming mutation-based resistance to antiandrogens with rational drug design. *Elife* **2**, e00499 (2013).

**Acknowledgements** We thank A. Paliakov, T. Barrette, Y. Qiao, P. Vats, R. Stender, X. Jiang, M. Pranihi and S. Han for technical assistance; C. Kumar-Sinha, M. Dhanasekaran, N. Palanisamy and P. Kunju for helpful discussions; J. Athanikar and K. Giles for critically reading the manuscript and submission of documents. This work was supported by a Movember-Prostate Cancer Foundation Challenge Award and in part by the Early Detection Research Network (U01 CA111275) and the NCI Prostate SPOR (P50CA69568) to A.M.C. A.M.C. is also supported by the Doris Duke Charitable Foundation, American Cancer Society, and A. Alfred-Taubman Institute. I.A.A. and R.A. are supported by a PCF Young Investigator Award.

**Author Contributions** I.A.A. and A.M.C. conceived the study and the experiments. I.A.A. performed the experiments with assistance from V.L.D., R.M., S.D., C.E. and X.J.; X.W. carried out *in vitro* interaction studies; M.C., R.M. and I.A.A. analysed microarray data; J.E.-W., K.W.-R. and F.Y.F. performed mouse xenograft studies; Y.-M.W. generated ChIP-seq libraries and X.C. performed the sequencing; M.C., R.Y., M.K.I. and Z.S.Q. performed ChIP-seq analysis with input from I.A.A.; S.W. provided compounds. I.A.A. and A.M.C. wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Gene expression and ChIP-seq coordinates have been deposited in the Gene Expression Omnibus under accession number GSE55064. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.C. ([arul@umich.edu](mailto:arul@umich.edu)).

## METHODS

**Cell culture.** VCaP prostate cancer cells were grown in DMEM with Glutamax (Gibco); LNCaP, 22RV1, DU145 and PC3 prostate cancer cell lines were grown in RPMI 1640; all were supplemented with 10% FBS (Invitrogen) in 5% CO<sub>2</sub> cell culture incubator. The immortalized benign prostate cell line RWPE-1 was grown in keratinocyte media with supplements (Lonza). All cell lines were tested and found to be free of mycoplasma contamination.

**Cell viability assay.** Cells were seeded in 96-well plates at 2,000–10,000 cells per well (optimum density for growth) in a total volume of 100 µl media containing 10% FBS. Serially diluted compounds in 100 µl media were added to the cells 12 h later. After 96 h incubation, cell viability was assessed by Cell-Titer GLO (Promega). The values were normalized and IC<sub>50</sub> was calculated using GraphPad Prism software. For long-term colony formation assay, 10,000–50,000 cells per well were seeded in 6-well plates and treated with either 100 nM or 500 nM of JQ1 or DMSO. After 12 days, cells were fixed with methanol, stained with crystal violet and photographed. For colorimetric assays, the stained wells were treated with 500 µl 10% acetic acid and the absorbance was measured at 560 nm using a spectrophotometer.

**Cell cycle analysis.** Cells were grown in 6-well plates and treated with varying concentrations of JQ1. For cell cycle analysis, cells were washed 48 h post-treatment with PBS and fixed in 70% ethanol overnight. The cells were washed again with PBS, stained with propidium iodide and analysed by flow cytometry.

**RNA interference.** For knockdown experiments, cells were seeded in 6-well plates and transfected with 100 nM ON-TARGETplus SMARTpool siRNA (ThermoScientific) targeting BRD2, BRD3, BRD4, MYC or non-targeting control (Non-targeting Pool, catalogue no. D-001810-10-50, using oligofectamine (Invitrogen) according to the manufacturer's instructions). The following are the catalogue numbers and the siRNA sequences: ON-TARGETplus Human BRD2 SMARTpool, catalogue no. L-004935-00-0005, target sequences, CACGAAAGCUACAGGAUGU, GGGCCGAGUUGUGCAUUA, CCUAAGAAGUCCAAGAAAG, GUCCUUCUGGCCUACGUA; ON-TARGETplus Human BRD3 SMARTpool, catalogue no. L-004936-00-0005, target sequences, AAUUGAACCGCCGGAUUA, CGGCUGAUGUUCUGAAUU, GGAGAGAUUGUCAGUCU, GCGAAUGUAUGCAGGACUU; ON-TARGETplus Human BRD4 SMARTpool, catalogue no. L-004937-00-0005, target sequences, AAACCGAGAUCAUGAUAGU, CUACACGACUACUGUGACA, AAACACAACUACAGCAUCG, CAGCGAAGACUCCGAAACA; and ON-TARGETplus Human MYC SMARTpool, catalogue no. L-003282-00-0005. Cells were trypsinized 24 h post-transfection and used in cell proliferation and Matrigel invasion assays as well as for RNA extractions to determine the knockdown efficiency.

For AR knockdown ON-TARGETplus Human AR SMART pool, catalogue no. L-003400-00-0005, target sequences, GAGCGUGGACUUCGGA, UCAAGGAACUCGAUCGUUA, CGAGAGAGCUGCAUCAGUU, CAGAAUGAUUGCAUAUU was used at 100 nM concentration; for ERG knockdown siRNA from Dharmacon, catalogue no. D-003886-01-0050 was used; and for ETV1 knockdown a mix of ETV1 siRNA ID s4854, catalogue no. 4392420 and ETV1 siRNA ID s4855, catalogue no. 4392420 from Life Technologies were used at 100 nM concentration for transfection using oligofectamine.

**Cell proliferation assay.** For cell proliferation assays after siRNA knockdown, 20,000 cells per well were seeded in 24-well plates ( $n = 3$ ) and cells were harvested and counted at the indicated time points by Coulter counter (Beckman Coulter).

VCaP, LNCaP and 22RV1 cells were transduced with either Ad-c-MYC (Vector Biolabs, catalogue no. 1285) or LacZ control adenoviral particles. Twenty-four hours after infection, equal number of cells were seeded in 24-well plates and treated with vehicle, JQ1 or I-BET762 at 500 nM concentration. Cells were counted at the indicated time points by Coulter Counter.

**Matrigel invasion assays.** Twenty-four hours after infection with siRNA or 500 nM JQ1 treatment,  $0.2 \times 10^6$  VCaP or  $0.1 \times 10^6$  LNCaP cells were seeded in a transwell chamber pre-coated with Matrigel (BD Biosciences). Medium containing 10% FBS in the lower chamber served as chemoattractant. In the case of JQ1, 500 nM compound was added to both upper and lower chambers. After 48 h, the non-invading cells and extracellular matrix were gently removed with a cotton swab and invasive cells located on the lower side of the chamber were stained with crystal violet, air dried, photographed and counted.

PC3 and RWPE cells were treated with JQ1 or I-BET762 at 500 nM concentration along with DMSO control for 24 h before seeding 50,000 cells per well in a transwell chamber pre-coated with Matrigel along with the corresponding drugs used for treatment. Medium containing 10% FBS in the lower chamber served as chemoattractant. After 48 h, the non-invading cells and extracellular matrix were gently removed with a cotton swab and invasive cells located on the lower side of the chamber were stained with crystal violet, air dried and photographed. For colorimetric assays, the inserts were treated with 150 µl of 10% acetic acid and the absorbance measured at 560 nm using a spectrophotometer (GE Healthcare).

**RNA isolation and quantitative real-time PCR.** Total RNA was isolated from cells using RNeasy Mini Kit (Qiagen) and cDNA was synthesized from 1,000 ng total RNA

using SuperScript III First-Strand Synthesis SuperMix (Invitrogen). qPCRs were performed in duplicate or triplicate using Taqman assays (Applied Biosystems) or standard SYBR green reagents and protocols on a StepOnePlus Real-Time PCR system (Applied Biosystems). The target mRNA expression was quantified using the  $\Delta\Delta C_t$  method and normalized to GAPDH expression. All primers were designed using primer 3 (<http://frodo.wi.mit.edu/primer3/>) and synthesized by Integrated DNA Technologies. The primer sequences for the SYBR green and catalogue numbers for TaqMan assays qPCR used are as follows: *BRD2* qPCR fwd, CTACGTAAGAAA CCCCCGAAG; *BRD2* qPCR rev, GCTTTTTCTCCAAAGCCAGTT; *BRD3* qPCR fwd, CCTCAGGGAGATGCTATCCA; *BRD3* qPCR rev, ATGTCGTGGTAGTCG TGCAG; *BRD4* qPCR fwd, AGCAGCAACAGCAATGTGAG; *BRD4* qPCR rev, G CTTGCACTTGCTCTTCC; *ERG* qPCR fwd, CGCAGAGTATCGTGCCAGC AGAT; *ERG* qPCR rev, CCATATTCTTCCAGCCCACTCC; *PSA(KLK3)* qPCR fwd, ACGCTGGACAGGGGGCAAAAG; *PSA(KLK3)* qPCR rev, GGGCAGGGCA CATGGTTCAC; *TMPSR2* qPCR fwd, CAGGAGTGTACGGGAATGTGATG GT; *TMPSR2* qPCR rev, GATTAGCCGTCTGCCCTCATTTGT; *FKBP5* qPCR fwd, TCTCATGTCTCCCCAGTTCC; *FKBP5* qPCR rev, TTCTGGCTTTCACGT CTGTG; *SLC45A3* qPCR fwd, TCGTGGCGAGGGGGCTGTA; *SLC45A3* qPCR rev, CATCCGAACGCCCTTCATCATAGTGT; *BMPRI1* qPCR fwd, CCACCAT TGTCCAGAAGACTC; *BMPRI1* qPCR rev, GCAACCCAGAGTCATCTCTT; *MYC* qPCR fwd, GCTCGTCTCAGAGAAGCTGG; *MYC* qPCR rev, GCTCAGATC CTCAGGTACAA; *AR* qPCR fwd, CAGTGGATGGGTGAAAAAT; *AR* qPCR rev, GGAGCTTGGTGAGCTGGTGA; *ETV1* qPCR fwd, GCAAGAAGGCTTCT GGCTCAT; *ETV1* qPCR rev, CCTTCCCGATACATTCTGGCT; *GAPDH* qPCR fwd, TGCACCACCAACTGCTTAGC; *GAPDH* qPCR rev, GGCATGGACTGTG GTCATGAG; *MYC* dis.enh ChIPPCR fwd, TGGCAACTTCTGCCTGTGTA; *MYC* dis.enh ChIPPCR rev CAGGCAGGGAGGAAGTCAAT; *MYC* upstream ChIPPCR fwd, CCAGGACAAATGACCACAGA; *MYC* upstream ChIPPCR rev, CCCTTGGCAAACATCAACTT; TaqMan primer probes *TDRD1*, catalogue no. Hs00229805 m1; *CACNA1D*, catalogue no. Hs00167753 m1; *ARHGDI1* catalogue no. Hs00171288 m1; *NDRG1*, catalogue no. Hs00608387 m1; *VCL*, catalogue no. Hs00419715 m1; *KRT8*, catalogue no. Hs01595539 g1; *MALAT1*, catalogue no. Hs00273907 s1; *BCL-xl* qPCR, catalogue no. Hs00236329 m1; *WNT2* qPCR, catalogue no. Hs00608224 m1; *CRISP3* qPCR, catalogue no. Hs00195988 m1.

**Antibodies and immunoblot analyses.** Antibodies used in the immunoprecipitation (IP) and immunoblotting (IB) assays are AR IP, IB (Abcam catalogue no. ab74272); RNA Pol II IB (Abcam catalogue no. ab5408); BRD2 IB (Abnova catalogue no. PAB3245); BRD3 IB (SantaCruz catalogue no. sc-81202); BRD4 IB (Bethyl catalogue no. A301-985A); ERG IB (Epitomics catalogue no. 2805-1); MYC IB (Sigma catalogue no. M5546); PSA IB (Dako catalogue no. A0562); GST IB (GE Life Science catalogue no. 27-4577-01); Halo IP, IB (Promega catalogue no. G9281); Poly Histidine IP, IB (Sigma catalogue no. H1029); BCL-xl IB (Cell Signaling catalogue no. 2762); cPARP IB (Cell Signaling catalogue no. 9541); GAPDH (14C10) IB (Cell Signaling catalogue no. 3683 s). All antibodies were used at dilutions suggested by the manufacturers. For western blot analysis, 200 µg total protein extract was boiled in sample buffer and 10–20 µg aliquots were separated by SDS-PAGE and transferred onto polyvinylidene difluoride membrane (GE Healthcare). The membrane was incubated for 1 h in blocking buffer (Tris-buffered saline, 0.1% Tween (TBS-T), 5% non-fat dry milk) followed by incubation overnight at 4 °C with the primary antibody. After a wash with TBS-T, the blot was incubated with horseradish peroxidase (HRP)-conjugated secondary antibody and signals were visualized by enhanced chemiluminescence system as per the manufacturer's protocol (GE Healthcare).

**Immunoprecipitations.** For endogenous immunoprecipitation experiments, nuclear extracts were obtained from VCaP and LNCaP cells using NE-PER nuclear extraction kit (Thermo Scientific). Nuclear pellet was then lysed in IP buffer (20 mM Tris pH 7.5, 150 mM NaCl, 1% Triton-X 100, protease inhibitor) by sonication. Nuclear lysates (0.5–1.0 mg) were pre-cleaned by incubation with protein G Dynabeads (Life Technologies) for 1 h on a rotator at 4 °C. Five micrograms of antibody was added to the pre-cleaned lysates and incubated on a rotator at 4 °C overnight before the addition of protein G Dynabeads for 1 h. Beads were washed three times in IP buffer and resuspended in 40 µl of 2× loading buffer and boiled at 90 °C for 10 min for separation of the protein and beads. Samples were then analysed by SDS-PAGE and western blotting as described earlier. For endogenous competitive assays, the VCaP cells were incubated with 5 or 25 µM JQ1 for 6 h before nuclear protein extractions.

For co-immunoprecipitation experiments in 293T cells, plasmids encoding various deletion mutants of BRD4 in pCDNA4c (Addgene) and full-length AR in pFN21 plasmid (Promega) were transfected using Fugene 6.0 HD (Roche) according to the manufacturer's instructions. Twenty-four hours after transfection, total proteins were extracted using IP buffer supplemented with protease inhibitor cocktail mix (Sigma) and the expressions of the corresponding proteins were analysed by immunoblotting. Immunoprecipitation using Halo beads followed by immunoblotting with anti-His antibody were performed as described earlier.



**Cell-free protein–protein interaction studies.** *In vitro* protein expression was carried out by cloning the desired expression cassettes downstream of a Halo or GST tag to produce fusion proteins. Briefly, AR and its subdomains were cloned into the pFN2K vector containing N-terminal GST sequence (catalogue no. G1891, Promega); BRD4 and its subdomains were cloned into the pFN19A vector containing N-terminal Halo sequence (catalogue no. C8461, Promega). After cloning, the fusion proteins were expressed using the cell-free transcription and translation system (catalogue no. L5030, Promega) following the manufacturer's protocol. For each reaction, protein expression was confirmed by western blot.

A total of 10 µl cell-free reaction containing Halo- and GST-tag fusion proteins were incubated in PBST (0.1% Tween) at 4 °C overnight. Ten microliters of HaloLink beads (catalogue no. G931, Promega) were blocked in BSA at 4 °C overnight. After washes with PBS, the beads were mixed with AR–BRD4 mixture and incubated at room temperature for 1 h. HaloLink beads were then washed with PBST four times and eluted in SDS loading buffer. Proteins were separated on SDS gel and blotted with anti-GST antibody (GE healthcare). For competitive assay, AR–BD1, NTD1b–BD1 and AR–BD2 mixture was incubated in the presence of varying doses of JQ1 compound.

**AR–BRD4 direct interaction assays by OctetRED.** The binding affinity between AR and BRD4 was determined by biolayer interferometry technology using the OctetRED system (ForteBio). Recombinant AR protein (catalogue no. AR-8486H, Creative Biomart) was biotinylated by EZ-Link NHS-PEG4 Biotinylation Kit (catalogue no. 21329, Thermo Scientific) following the manufacturer's protocol, and any unincorporated biotin was removed from the reactions with Zeba 2 ml desalt columns. Biotinylated proteins (5 µg ml<sup>−1</sup>) were then incubated with super-streptavidin biosensors (catalogue no. 18-5057, ForteBio) in binding buffer (20 mM HEPES pH 7.4, 150 mM NaCl) and washed three times in binding buffer. BRD4 (BD1–BD2) protein (catalogue no. 31047, BPS Biosciences) was serially diluted in binding buffer, and the AR–BRD4 association/dissociation was monitored by OctetRED for 10 min at 25 °C. Non-specific binding was controlled by subtracting the signal obtained from AR–RNF2 interactions from that of AR–BRD4 interactions and baseline signal drift was controlled by monitoring immobilized AR without BRD4. OctetRED analysis software was used to analyse the data.

**Gene expression array analysis.** VCaP, LNCaP, 22RV1 and DU145 cells were treated with 500 nM JQ1 for 24 h and total RNA was extracted using RNeasy Mini Kit (Qiagen) for gene expression array analysis. For anti-androgen comparative study, VCaP and LNCaP cells were grown in media containing 10% charcoal-stripped serum for 48 h followed by pre-treatment with 500 nM JQ1, 10 µM MDV3100 or 25 µM bicalutamide for 6 h and stimulated with 10 nM DHT (androgen) for 18 h. Cells treated with only vehicle or 10 nM DHT served as controls. For determining the effect of BET inhibitors in isogenic ERG system, RWPE-ERG and PC3-ERG cells were treated with 500 nM JQ1 or I-BET762 for 24 h. Expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray according to the manufacturer's protocol. All samples were run in technical duplicates or quadruplets against control. Over- and underexpressed gene sets were generated by filtering to include only data points that showed twofold average over- or underexpression (log ratio with  $P < 0.001$ ) in all hybridizations.

GSEA was performed using the JAVA program (<http://www.broadinstitute.org/gsea>) as described previously<sup>31</sup>.

The AR target gene signature used in GSEA analysis was generated from common upregulated genes in VCaP and LNCaP upon DHT treatment and the gene list is as follows: *ABCC4*, *ABHD2*, *ACSL3*, *ADARB2*, *AF349445*, *AFF4*, *AI089002*, *AI207522*, *AI570240*, *AK023660*, *AK025360*, *AK055915*, *AK057576*, *AK074291*, *AK092594*, *AK093002*, *AK098478*, *AK124281*, *AK124426*, *AL533190*, *AL713762*, *ALDH1A3*, *AMACIL2*, *ANKRD37*, *ANXA2*, *ARSG*, *ASRGL1*, *ATP10A*, *ATP1A1*, *ATP1A4*, *ATRNL1*, *AUTS2*, *AW029229*, *AW389914*, *AZGP1*, *B3GAT1*, *BC039021*, *BC041926*, *BC041955*, *BC055421*, *BC062780*, *BG462058*, *BG618474*, *BI710972*, *BM469851*, *BMPRI3*, *BQ017638*, *BQ706262*, *BRP44*, *BU567141*, *BU753102*, *BX099483*, *C10orf114*, *C14orf162*, *C16orf30*, *C18orf1*, *C1orf108*, *C1orf113*, *C1orf26*, *C2orf112*, *C6orf81*, *CA314451*, *CA414006*, *CBL1*, *CCDC4*, *CDC14B*, *CDC14C*, *CDYL2*, *CEBPD*, *CENPN*, *ChGn*, *CHIA*, *CHKA*, *CHST2*, *CLDN12*, *CLDN14*, *CLDN8*, *CTBP1*, *CUTL2*, *CXorf9*, *CYP1A1*, *CYP2U1*, *DDR2*, *DHCR24*, *DKFZp761P0423*, *DNAJB9*, *DOCK8*, *EAF2*, *EDG7*, *ELL2*, *ELOVL5*, *ELOVL7*, *EMPI1*, *ENDOD1*, *ENST00000358356*, *ERN1*, *ERRFI1*, *F2RL1*, *FAM13A10S*, *FER1L3*, *FGD4*, *FKBP5*, *FLJ31568*, *FLJ39502*, *FRK*, *FZD5*, *GADD45G*, *GIPR*, *GREB1*, *GSR*, *HERC3*, *HLA-DRB3*, *HOMER2*, *HPGD*, *HS3ST4*, *HSD17B2*, *IFI6*, *IGF1*, *IGFIR*, *IL20RA*, *IMPAD1*, *INPP4B*, *KCNMA1*, *KLF15*, *KLK3*, *KLK4*, *KLK5*, *KRT18*, *KRT19*, *KRT72*, *LAMA1*, *LDLR*, *LIFR*, *LOC205251*, *LOC401708*, *LOC641467*, *LOC646282*, *LOC730498*, *LONRF1*, *LOX*, *LRCH1*, *LRIG1*, *LSS*, *MAF*, *MAK*, *MALT1*, *MAP1B*, *MAP7D1*, *MBOAT2*, *MFS2D*, *MICAL1*, *MLPH*, *MOGAT2*, *MPZL1*, *MTMR9*, *NANOGP1*, *NAT1*, *NCAPD3*, *NDFIP2*, *NDRG1*, *NEBL*, *NEK10*, *NFKBIA*, *NNMT*, *NR4A1*, *NY-REN-7*, *ODC1*, *OLAH*, *ORM1*, *ORM2*, *OTUD7B*, *PACSI1*, *PDLIM5*, *PECI*, *PER1*, *PFKFB2*, *PGC*, *PHACTR3*, *PNPLA8*, *PPP2CB*,

*RAB27A*, *RAB4A*, *RASD1*, *RHOA*, *RUNX1*, *S100A5*, *SCRG1*, *SGK*, *SHROOM3*, *SLC16A6*, *SLC26A2*, *SLC26A3*, *SLC2A14*, *SLC2A3*, *SLC38A4*, *SLC41A1*, *SLC45A3*, *SLITRK6*, *SMC4*, *SMOC1*, *SNAI2*, *SNHG2*, *SNTG2*, *SOC2*, *SPDEF*, *SPDYA*, *SPINK5L3*, *SPOCK1*, *SPTB*, *ST6GALNAC1*, *STEAP4*, *STK17B*, *TACCI*, *TBRG1*, *TBX15*, *TG*, *TGFB2*, *TIPARP*, *TLOC1*, *TMCC3*, *TMPPRS2*, *TNFAIP3*, *TPD52*, *TRIM36*, *TRIM63*, *TTN*, *TUBA3D*, *WIPI1*, *WNT7B*, *WWTR1*, *XO3757*, *ZBTB1*, *ZBTB16* and *ZBTB24*.

The ERG gene signature was generated by extracting twofold upregulated genes from RWPE and PC3 cells stably expressing ERG compared with respective LacZ-expressing cells. GSEA was performed using this gene set on gene expression data obtained from the JQ1- and I-BET762-treated RWPE and PC3 cells. We also ran GSEA using a gene set that was not changed upon expression of ERG to exclude the possibility that treatment with JQ1 and I-BET762 may change gene expression in a non-specific fashion. All of the gene expression array data (total 48) can be found at the Gene Expression Omnibus under accession number GSE55064.

**ChIP and ChIP-seq.** The ChIP assays for BRD2, BRD3, BRD4, AR, RNA Pol II, ERG and H3K27ac were performed using HighCell ChIP kit (Diagenode) according to the manufacturer's protocol. The antibodies used for ChIP assay are AR PG-21 (Millipore catalogue no. 06-680); RNA Pol II (Abcam catalogue no. ab5408); BRD2 (Bethyl catalogue no. A302-583A); BRD3 (Bethyl catalogue no. A302-368A); BRD4 (Bethyl catalogue no. A301-985A); H3(acetyl K27) (Abcam catalogue no. ab4729) and IgG (Diagenode). For BRD2/3/4 ChIP-seq experiments with BET inhibitors, VCaP cells were treated with 500 nM JQ1 or I-BET762 for 12 h. For AR signalling ChIP-seq experiments, VCaP cells were grown in charcoal-stripped serum containing media for 48 h followed by 6 h pre-treatment with vehicle or 500 nM JQ1 or 10 µM MDV3100 or 25 µM bicalutamide and then stimulated with 10 nM DHT for 12 h. For ERG ChIP-seq studies, VCaP cells were treated with 500 nM JQ1 or vehicle for 12 h. Next, cells were crosslinked for 10 min with 1% formaldehyde. Crosslinking was terminated by the addition of 1/10 volume 1.25 M glycine for 5 min at room temperature followed by cell lysis and sonication (Bioruptor, Diagenode), resulting in an average chromatin fragment size of 200 bp. Chromatin equivalent to 5 × 10<sup>6</sup> cells were used for ChIP using various antibodies. ChIP DNA was isolated (Pure Kit, Diagenode) from samples by incubation with the antibody at 4 °C overnight followed by washing and reversal of crosslinking. The ChIP-seq sample preparation for sequencing was performed according to the manufacturer's instructions (Illumina). ChIP-enriched DNA samples (1–10 ng) were converted to blunt-ended fragments using T4 DNA polymerase, *E. coli* DNA polymerase I large fragment (Klenow polymerase) and T4 polynucleotide kinase (New England Biolabs (NEB)). A single A base was added to fragment ends by Klenow fragment (3' to 5' exo minus; NEB) followed by ligation of Illumina adaptors (Quick ligase, NEB). The adaptor-modified DNA fragments were enriched by PCR using the Illumina Barcode primers and Phusion DNA polymerase (NEB). PCR products were size selected using 3% NuSieve agarose gels (Lonza) followed by gel extraction using QIAEX II reagents (Qiagen). Libraries were quantified with the Bioanalyzer 2100 (Agilent) and sequenced on the Illumina HiSeq 2000 Sequencer (100-nucleotide read length).

#### ChIP-seq analysis

**ChIP-seq enrichment levels.** ChIP enrichment levels within a peak (or site) were calculated from the sequencing data as follows. First, reads were aligned to the HG19 reference genome using Bowtie2<sup>32</sup> with all default settings. Second, aligned reads were sorted using NovoSort and exact duplicates were removed using Samtools<sup>33</sup>. Third, for each peak (site) overlapping reads were counted and this count was divided by the length of the peak or site. Fourth, to correct for differences in sequencing depth and alignment coverage the values are further normalized by the number of aligned reads per million.

**ChIP-seq reproducibility plots.** To assess the biological variability of AR and ERG ChIP-seq experiments, we compared enrichment levels of their respective replicates. For each replicate we called peaks using MACS with all default setting against an IgG control. We excluded peaks within genomic regions prone to technical artefacts<sup>34</sup>. For each replicate pair we defined a set of concordant peaks as those overlapping in both replicates. For each concordant peak we calculated enrichment levels within the union of the two overlapping peaks. The scatter plots include all peaks with enrichment levels up to the 99th percentile.

**Overlaps of bromodomain proteins.** We compared the genome-wide distribution of BRD2, BRD3 and BRD4 peaks in DMSO-treated VCaP cells. First, we called peaks for each of the proteins using MACS with all default settings and IgG control. Because we were interested in peaks that are possibly biologically significant we used a moderately stringent significance cut-off (MACS score >100). Next, we identified all genomic regions that were enriched for at least one of the proteins. Specifically, we 'reduced' all stringent peaks using GenomicRanges<sup>35</sup>. For each of those regions we established which of the bromodomain proteins were enriched to count the number of overlaps.

**Drug-induced changes of bromodomain protein enrichment levels.** For each protein (BRD2, BRD3, BRD4) we assessed quantitative changes in their respective

enrichment levels upon drug treatment (I-BET762, JQ1) relative to the levels in the DMSO control. First, peaks were called for all conditions and proteins as described earlier. Next, for each protein separately, we identified genomic regions that were enriched in any (union) of the treatment conditions (DMSO, I-BET762 or JQ1). Within those regions we quantified enrichment levels as described earlier for determining ChIP-seq enrichment levels. As enrichment levels of different proteins are not directly comparable, we normalize all enrichments to the median level of the DMSO control.

**Differential AR–BRD4 enrichment and AR–BRD4 overlap.** HPeak, a hidden Markov model (HMM)-based peak-calling software<sup>36</sup> designed for the identification of protein-interactive genomic regions, was used for ChIP-seq peak determination. For enrichment plots shown in Fig. 3a, c and d, identified peaks for each sample are centred by peak summit and average coverage per million was counted within 1,500 bp relative to the peak centre. The overlap of AR- and BRD4-enriched regions were calculated by BEDtools<sup>37</sup>. The significance of overlap between AR and BRD4 binding was calculated using a hypergeometric test based on the derived number of associated genes. The heatmap for AR peak enrichment was generated using python-based script on raw data and visualized using JavaTreeView<sup>38</sup>.

**Differential ERG enrichment.** We identified sites with significant differences in ERG levels between DMSO- and JQ1-treated cells. First we focused on concordant peaks (see ChIP-seq reproducibility plots) that were overlapping or in the  $\pm 5$  kb proximity of annotated gene loci. We defined a gene locus as the union of all of its known transcripts (Ensembl Genes 73). We used DESeq2 to assess the statistical significance of differences in ERG enrichment levels. Although DESeq2 was originally developed for RNA-seq, its statistical model is well-suited to count data in general. We used the tools' default multiple hypothesis correction method and report peaks with significant differences in ERG levels (adjusted  $P$  value  $< 0.1$ ). To assess quantitative differences in ERG levels at significantly 'gained' (positive difference in ERG levels upon JQ1 treatment) and 'lost' (negative difference in ERG levels upon JQ1 treatment) we followed the same procedure as described earlier for determining ChIP-seq enrichment levels.

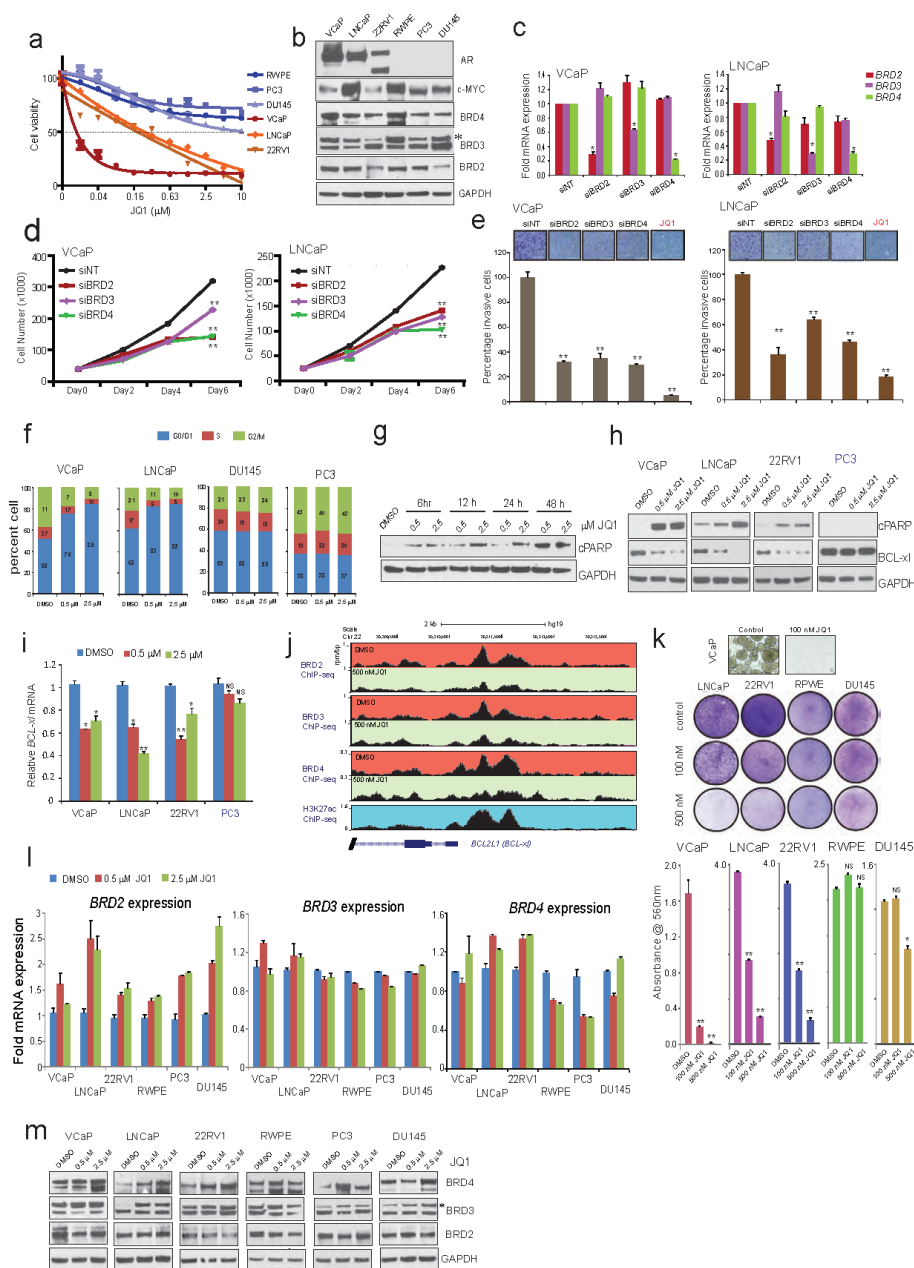
**Murine prostate tumour xenograft model.** Four-week-old male SCID C.B17 mice were procured from a breeding colony at University of Michigan maintained by our group. Mice were anaesthetized using 2% isoflurane (inhalation) and  $2 \times 10^6$  VCaP prostate cancer cells suspended in 100  $\mu$ l of PBS with 50% Matrigel (BD Biosciences) were implanted subcutaneously into the dorsal flank on both sides of the mice. Once the tumours reached a palpable stage (100 mm<sup>3</sup>), the animals were randomized and treated with either 10 mg kg<sup>-1</sup> body weight MDV3100 or 50 mg kg<sup>-1</sup> body weight (doses previously used in mouse prostate cancer and multiple myeloma models<sup>11,39</sup>) by oral gavage or intraperitoneally, respectively, for five days a week. Growth in tumour volume was recorded using digital callipers and tumour volumes were estimated using the formula ( $\pi/6$ ) ( $L \times W^2$ ), where  $L$  is length of tumour and  $W$  is width. Loss of body weight during the course of the study was also monitored. At the end of the studies mice were killed and tumours extracted and weighed. Additionally, femur bone marrow, liver and spleen were harvested to determine spontaneous metastasis by measuring

human-Alu sequence. Briefly, genomic DNA from femur bone marrow, liver and spleen were prepared using Puregene DNA purification system (Qiagen), followed by quantification of human Alu sequence by human Alu-specific Fluorogenic TaqMan qPCR probes as described previously<sup>40,41</sup>. For CRPC experiment, VCaP tumour-bearing mice were castrated when the tumours were approximately 200 mm<sup>3</sup> in size and once the tumour grew back to the pre-castration size were randomized and treated with JQ1 or vehicle (D5W) control. All procedures involving mice were approved by the University Committee on Use and Care of Animals at the University of Michigan and conform to all regulatory standards.

**Prostate histology and hormone measurement.** Four-to-five-weeks-old male SCID C.B17 mice were administered vehicle, 10 mg kg<sup>-1</sup> MDV3100 or 50 mg kg<sup>-1</sup> JQ1 by oral gavage or intraperitoneally, respectively, for five days a week. Highly hormone-responsive seminal vesicles attached to the prostate were harvested from mice after 4 weeks of injection. Prostate were fixed in formalin solution and processed for sectioning. Standard haematoxylin and eosin staining was performed on the formalin-fixed sections, which were used to image the different lobes of the gland. To determine testosterone levels, blood samples were collected by cardiac puncture from mice anaesthetized with isoflurane. The serum was separated from the blood and stored at  $-80^\circ\text{C}$  until assayed. Serum testosterone levels were measured by ligand assay at the University of Michigan-ULAM Pathology Cores for Animal Research.

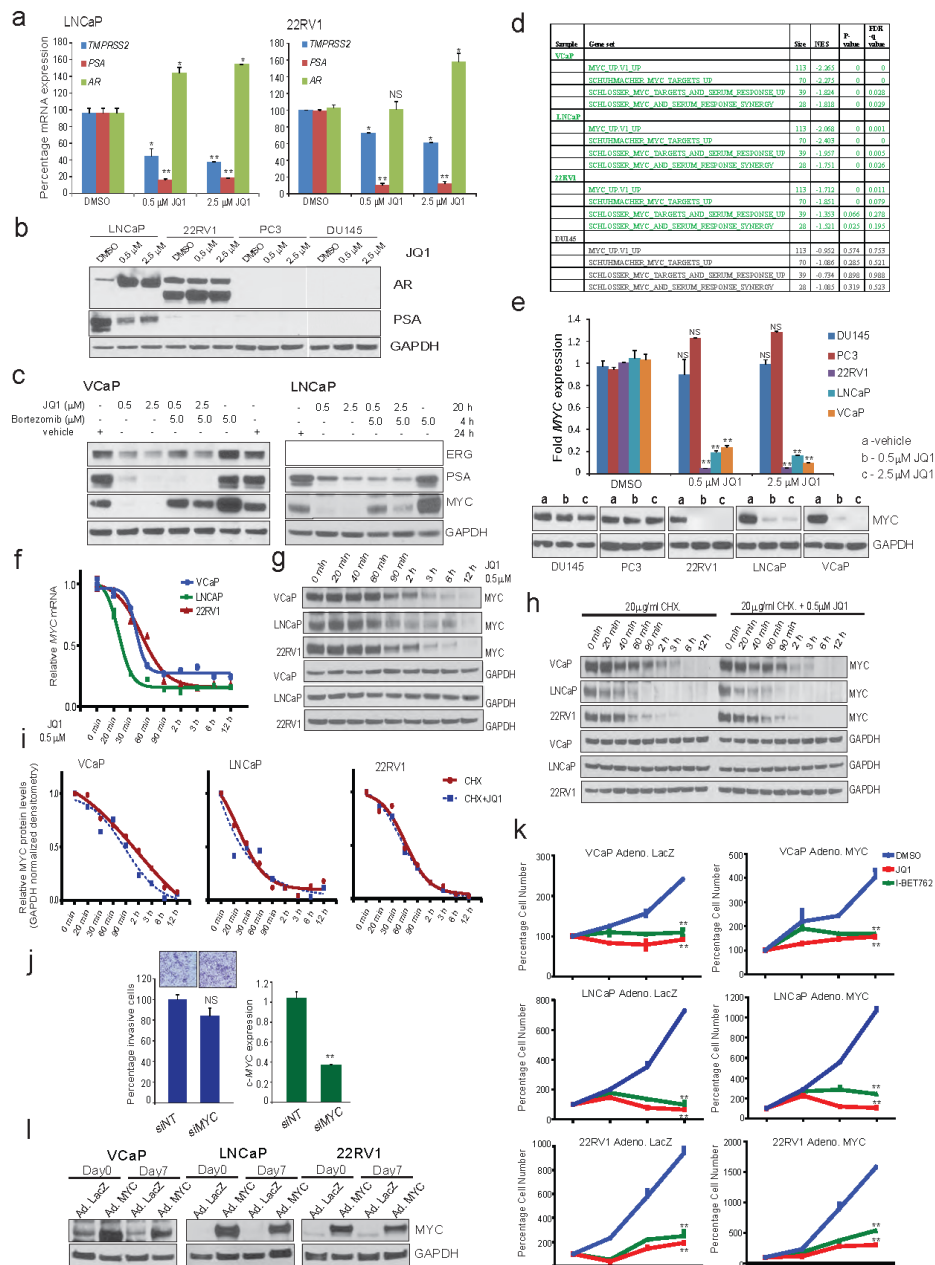
31. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
32. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
33. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
34. Pickrell, J. K., Gaffney, D. J., Gilad, Y. & Pritchard, J. K. False positive peaks in ChIP-seq and other sequencing-based functional assays caused by unannotated high copy number regions. *Bioinformatics* **27**, 2144–2146 (2011).
35. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
36. Qin, Z. S. *et al.* HPeak: an HMM-based algorithm for defining read-enriched regions in ChIP-Seq data. *BMC Bioinformatics* **11**, 369 (2010).
37. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
38. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
39. Tran, C. *et al.* Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* **324**, 787–790 (2009).
40. Asangani, I. A. *et al.* Characterization of the EZH2-MMSET histone methyltransferase regulatory axis in cancer. *Mol. Cell* **49**, 80–93 (2013).
41. van der Horst, E. H., Leupold, J. H., Schubert, R., Ullrich, A. & Allgayer, H. TagMan-based quantification of invasive cells in the chick embryo metastasis assay. *Biotechniques* **37**, 940–942, 944, 946 (2004).
42. Gao, L. *et al.* Androgen receptor promotes ligand-independent prostate cancer progression through c-Myc upregulation. *PLoS ONE* **8**, e63563 (2013).





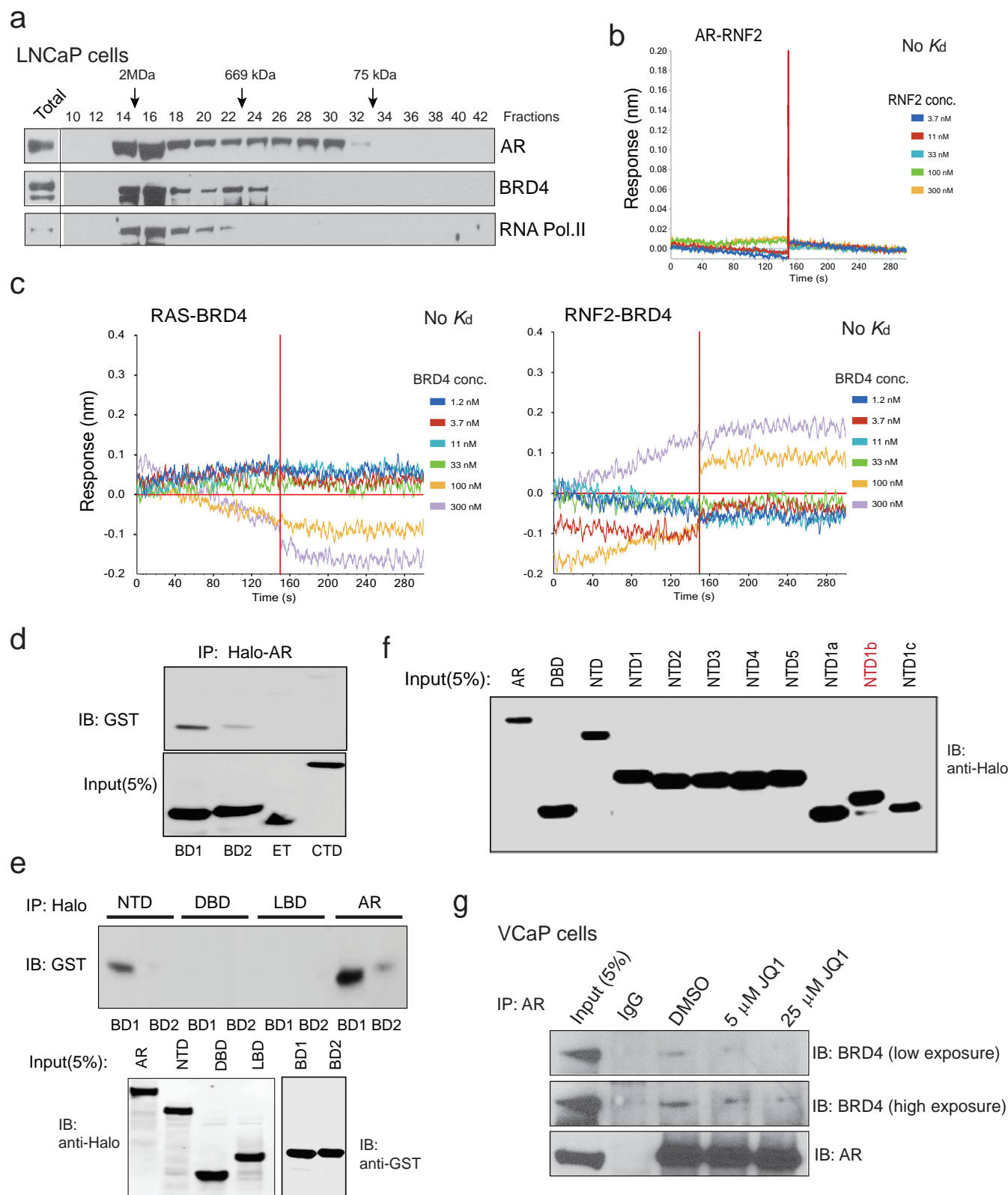
**Extended Data Figure 1 | BET bromodomain inhibitor JQ1 blocks cell growth, induces apoptosis and transcriptionally suppresses anti-apoptotic factor BCL-xl without affecting BRD2/3/4 proteins.** **a**, Cell viability curves for the six prostate lines treated with JQ1.  $N = 6$  wells of a 96-well plate per condition. **b**, BET bromodomain proteins are ubiquitously expressed in prostate cell lines. AR and MYC protein levels are also shown. GAPDH serves as a loading control. **c**, Knockdown of BET bromodomain proteins attenuates cell proliferation and invasion. qRT-PCR analyses of BRD2, BRD3 or BRD4 in VCaP cells transfected with short interfering RNA (siRNA) against their respective transcript or non-targeting (NT) siRNA. Data show mean  $\pm$  s.e.m. ( $n = 3$ ) from one of three independent experiments. **d**, VCaP- and LNCaP-cell proliferation after indicated gene knockdown. 20,000 cells were seeded in 24-well plates 24 h post-transfection with siRNAs and counted on day 0, 2, 4 and 6 ( $n = 3$ ) by Coulter counter. Data show mean  $\pm$  s.e.m. **e**, VCaP- and LNCaP-cell invasion ( $n = 6$ ) after indicated gene knockdown. JQ1 was used at 500 nM. **f**, Cell cycle analysis of JQ1-treated prostate cell lines (after 48 h treatment with JQ1). Data represent three independent experiments. **g**, Induction of apoptosis as determined by appearance of cleaved PARP (cPARP) in VCaP prostate cancer cells by JQ1. GAPDH served as a loading control. **h**, Immunoblot demonstrating an increase in cPARP and decrease in BCL-xl in all three AR-positive cell lines compared with AR-negative PC3 cells

upon JQ1 treatment. **i**, Relative *BCL-xl* mRNA levels as determined by TaqMan qPCR in JQ1-treated cells. Data show mean  $\pm$  s.e.m. ( $n = 3$ ) from one of three independent experiments. **j**, ChIP-seq data depicting loss of BRD2/3/4 recruitment to the BCL-xl promoter upon JQ1 treatment in VCaP cells. The genome browser representation of BRD2/3/4 binding events on the BCL-xl promoter region. The y-axis denotes reads per million per base pair (r.p.m. bp<sup>-1</sup>), the x-axis denotes the genomic position. The bottom panel depicts the H3K27ac mark on the same promoter region in VCaP cells. **k**, Colony formation assays of prostate cell lines. Cells were cultured in the presence or absence of 100 and 500 nM of JQ1 for 12 days followed by staining (top) and quantification (bottom; mean  $\pm$  s.e.m.  $n = 6$ ). Representative photographs of crystal violet stained colonies (except for VCaP) used for quantification are shown. **l**, BET bromodomain inhibitor JQ1 does not affect its target proteins. qRT-PCR analyses of BRD2, BRD3 and BRD4 in prostate cancer cell line panel treated with two different concentrations of JQ1 for 24 h. Data show mean  $\pm$  s.e.m. ( $n = 3$ ) from one of the three independent experiments. **m**, Immunoblot analysis of BRD proteins in prostate cell line panel treated with JQ1 for 48 hrs. GAPDH serves as a loading control. Asterisks in **b** and **m** indicate non-specific band. Representative blots shown are from triplicate biological experiment. NS, not significant; \* $P \leq 0.01$ ; \*\* $P \leq 0.001$  by two-tailed Student's *t*-test.



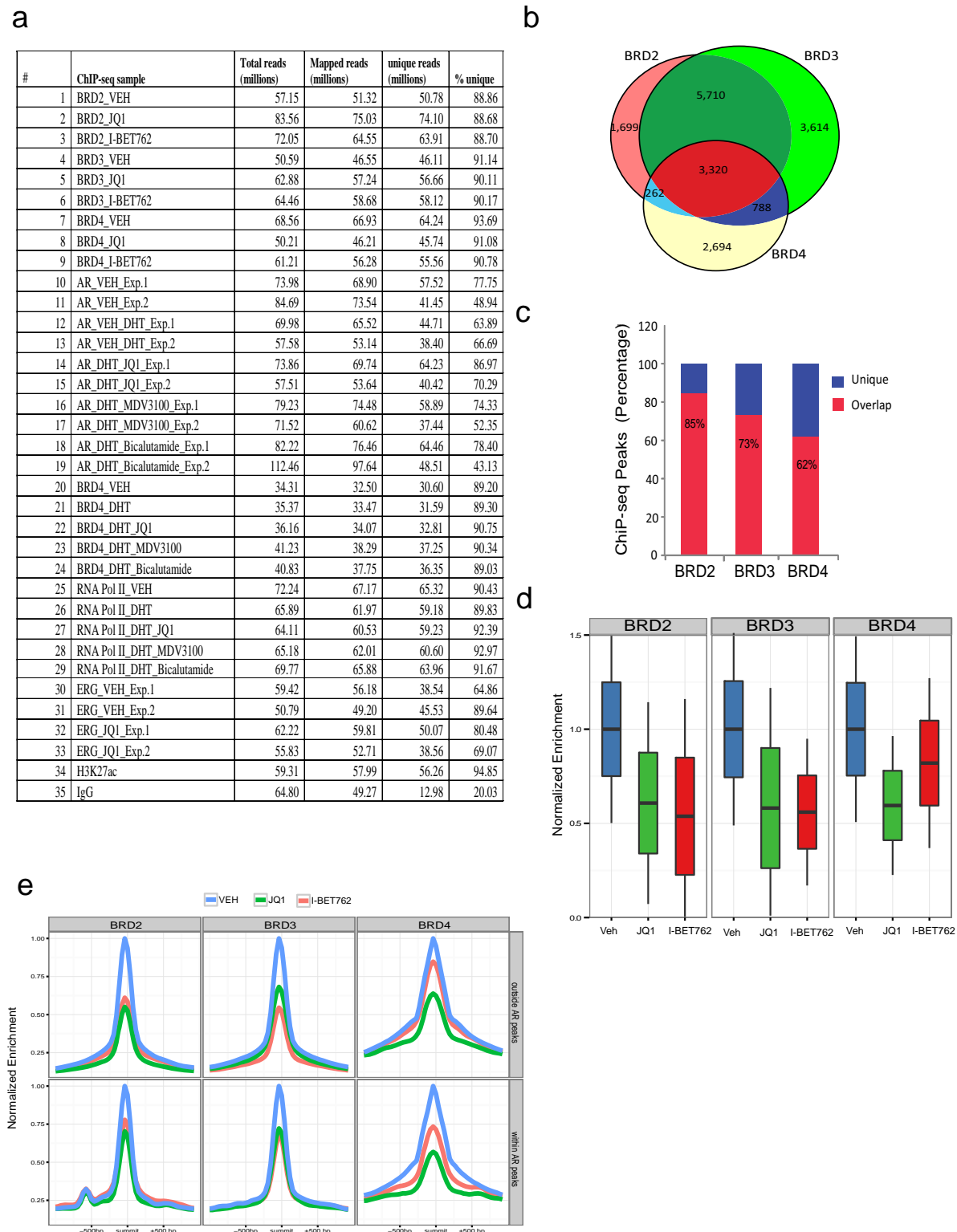
**Extended Data Figure 2 | Effect of JQ1 on AR target genes and on MYC transcription.** **a**, qRT-PCR analysis of indicated genes in LNCaP and 22RV1 cells treated with varying concentrations of JQ1 for 24 h. Data show mean  $\pm$  s.e.m. (n = 3) from one of two independent experiments. **b**, Immunoblot analysis of AR and PSA in a panel of prostate cancer cells after treatment with two different doses of JQ1. GAPDH serves as a loading control. **c**, ERG and PSA are transcriptional targets of JQ1. Proteasome inhibitor bortezomib does not rescue ERG and PSA levels in JQ1-treated VCaP cells. Immunoblot analyses of ERG and PSA in VCaP and LNCaP cells treated with JQ1 followed by incubation with bortezomib as indicated. MYC, known to be degraded by proteasome, was used as a positive control for bortezomib treatment. GAPDH serves as a loading control. **d**, GSEA showing loss of MYC signature (four-gene set) in AR-positive VCaP, LNCaP and 22RV1 cells but not AR-negative DU145 cells after JQ1 treatment. Size, number of genes in each set; NES, normalized enrichment score. P and false discovery rate (FDR) q values indicate statistical significance. **e**, qRT-PCR and immunoblot analysis of MYC in JQ1-treated prostate cancer cells. Data show mean  $\pm$  s.e.m. (n = 3) from one of two independent experiments. **f, g**, Time-course qRT-PCR and immunoblot analysis of MYC in AR-positive VCaP, LNCaP and 22RV1 cells after JQ1 treatment. **h**, Cycloheximide

(translation inhibitor) treatment does not enhance JQ1-mediated loss of MYC protein, ruling out post-translational degradation of MYC by JQ1. Time-course immunoblot analysis of MYC in VCaP, LNCaP and 22RV1 cells treated with cycloheximide or cycloheximide plus JQ1 as indicated. Representative blots from two independent experiments are shown. **i**, GAPDH-normalized MYC protein levels are shown. Band intensities from **d** were determined by ImageJ and the plots were generated using GraphPad Prism. **j**, MYC knockdown does not affect cell invasion. Box plot shows invasion of VCaP cells transfected with siNT or siMYC. Inset shows the image of invaded VCaP cells (n = 6). Right, qRT-PCR of MYC upon siRNA transfection. Data show mean  $\pm$  s.e.m. from one of three independent experiments. **k**, Exogenous MYC introduction does not rescue JQ1-mediated cell growth inhibition. Cells were infected with control adeno-LacZ or adeno-MYC virus. Equal numbers of cells were plated 24 h after infection and treated with 500 nM JQ1 or I-BET762. Cells were counted (n = 3 wells) and plotted; day 0 of drug treatment was set at 100%. Data show mean  $\pm$  s.e.m. from one of four independent experiments. **l**, Immunoblot analysis depicts overexpression of MYC in adeno-MYC infected cells on day 0 and day 7 of the experiment. GAPDH serves as a loading control. \*P  $\leq$  0.05; \*\*P  $\leq$  0.005 by two-tailed Student's t-test.



**Extended Data Figure 3 | Physical association of AR with BRD4 and its disruption by BET bromodomain inhibitor.** **a**, LNCaP nuclear extract was fractionated on a Superose-6 column and AR, BRD4 and RNA Pol II were analysed by immunoblot analysis. **b**, **c**, Representative sensorgrams for AR-RNF2, RAS-BRD4(BD1-BD2) and RNF2-BRD4(BD1-BD2) interactions by an OctetRED biolayer interferometry. Real-time binding was measured by immobilizing biotinylated AR, RAS or RNF2 proteins separately on a streptavidin biosensor and subsequent interaction with varying

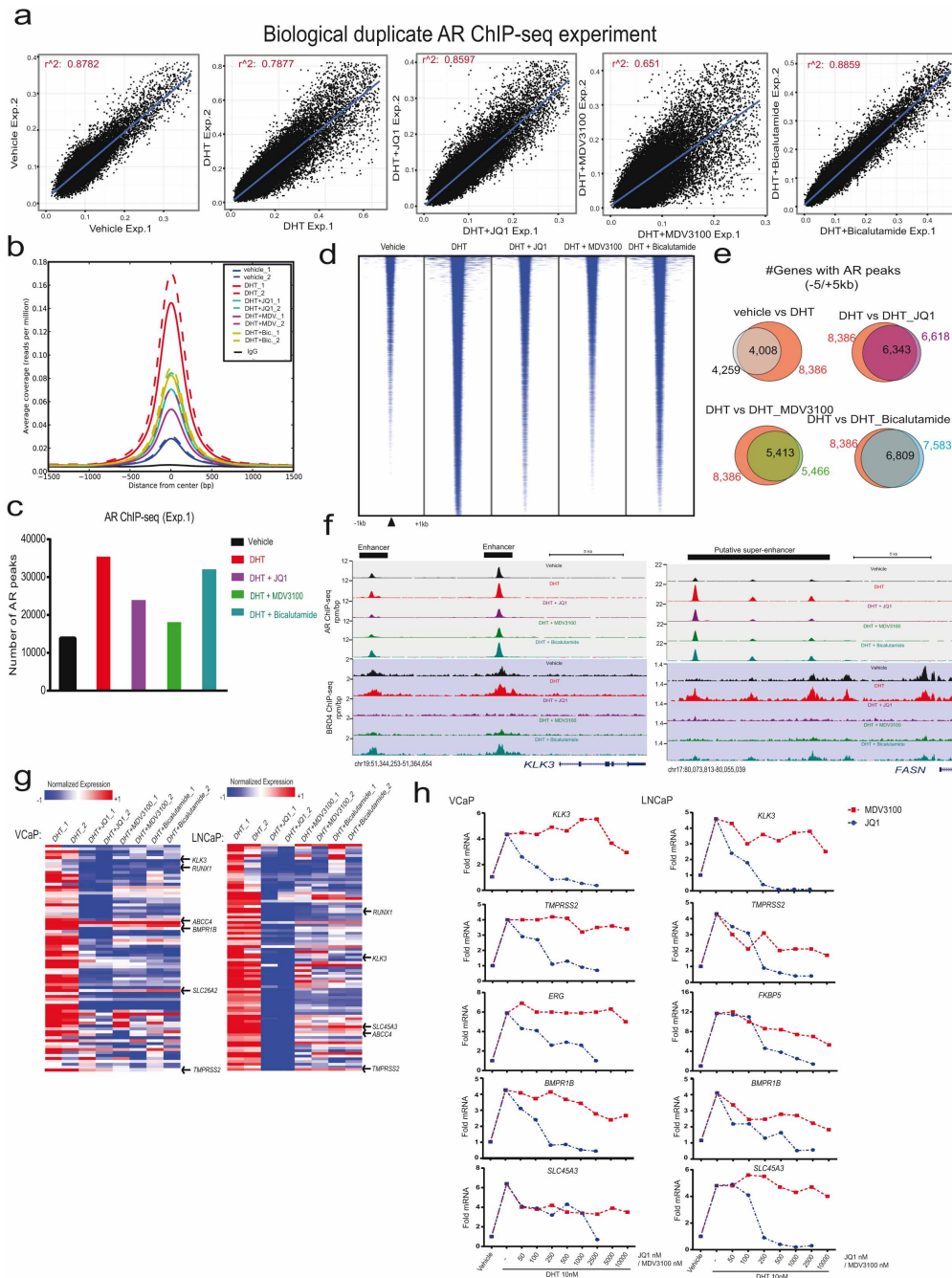
concentrations of analyte proteins (RNF2 or BRD4(BD1-BD2)) individually. Immobilized RAS or RNF2 biosensors did not bind with BRD4, indicating that the AR-BRD4 interaction is specific. Representative sensorgrams from 4-6 independent experiment are shown. **d-f**, *In vitro* binding analysis of AR and indicated domains of BRD4. Equal amounts of *in vitro* translated full-length Halo-AR protein and GST-BRD4 domains were combined and immunoprecipitated using Halo beads followed by immunoblot analysis with anti-GST antibody. **g**, JQ1 disrupts the endogenous AR-BRD4 interaction.



**Extended Data Figure 4 | Changes in genome-wide enrichment profiles of BRD proteins in response to bromodomain inhibitors.** **a**, Table showing high-throughput sequencing read information for ChIP libraries of BRD2, BRD3, BRD4, AR, RNA Pol II, ERG, H3K27ac and IgG performed for this study. **b**, ChIP-seq was performed using BRD2, BRD3 and BRD4 antibodies in VCaP cells treated with DMSO, JQ1 or I-BET762 for 12 h. Genome-wide distribution of BRD2, BRD3 and BRD4 enriched sites. Highly significant peaks (see Methods) show relatively high overlap. A large majority of sites are occupied by at least two BRD proteins. BRD2 and BRD3 have the most similar localization pattern. **c**, BRD proteins show varying degrees of overlap. Shown is

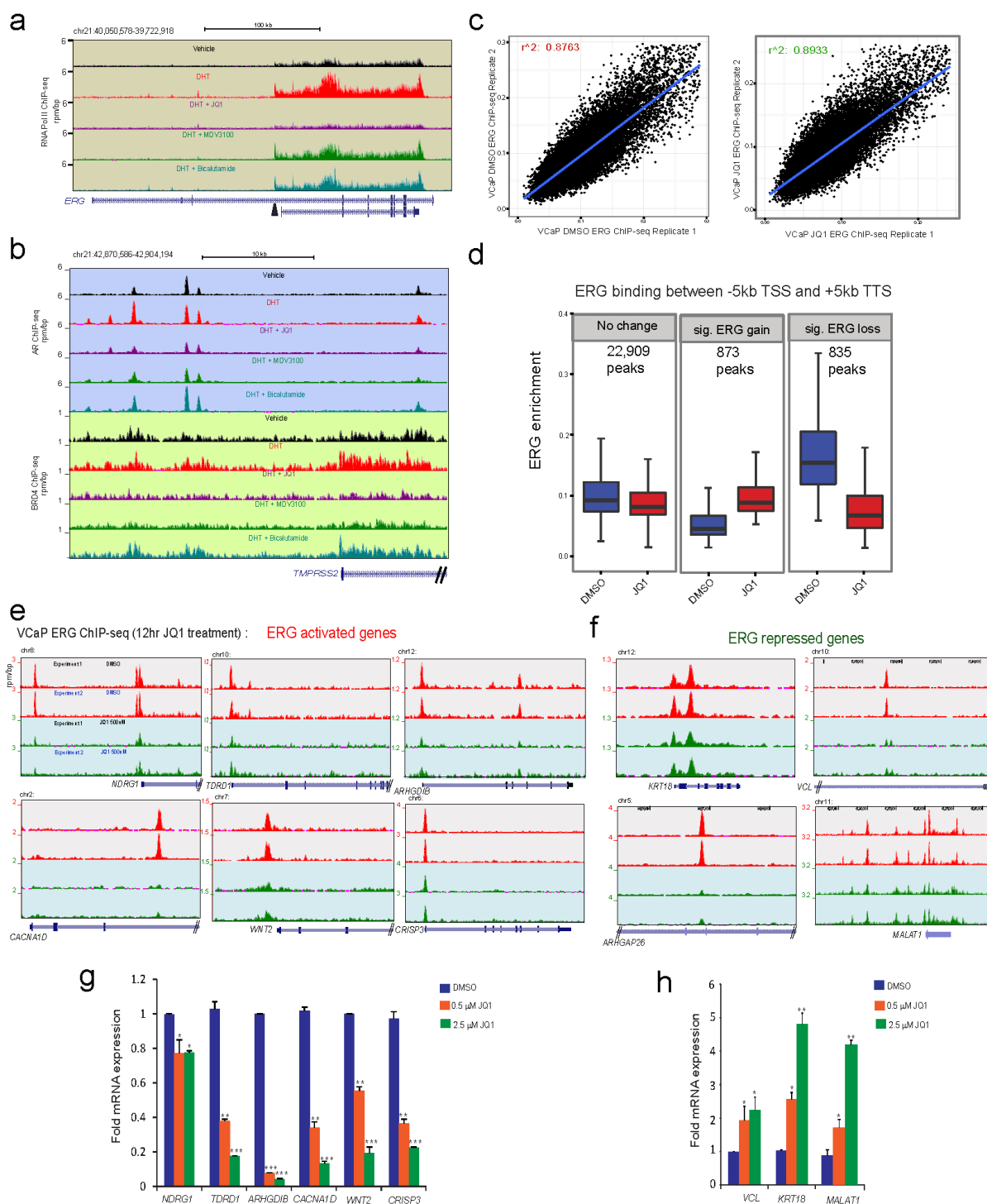
the ratio of sites occupied by either protein alone (unique) or co-occupied with another BRD-family protein (overlap). BRD4 shows the largest number of unique peaks. **d**, BET inhibitors JQ1 and I-BET762 attenuate recruitment of BRD proteins from chromatin. Enrichment levels for each protein were normalized to the median enrichment in vehicle-treated cells. BRD2 and BRD3 proteins show similar responses to both inhibitors, whereas BRD4 is more potently evicted by JQ1. **e**, BET bromodomain inhibitors deplete target proteins from genomic regions with or without AR. Mean enrichment levels within each subpanel were normalized to the maximum mean enrichment in vehicle-treated cells.





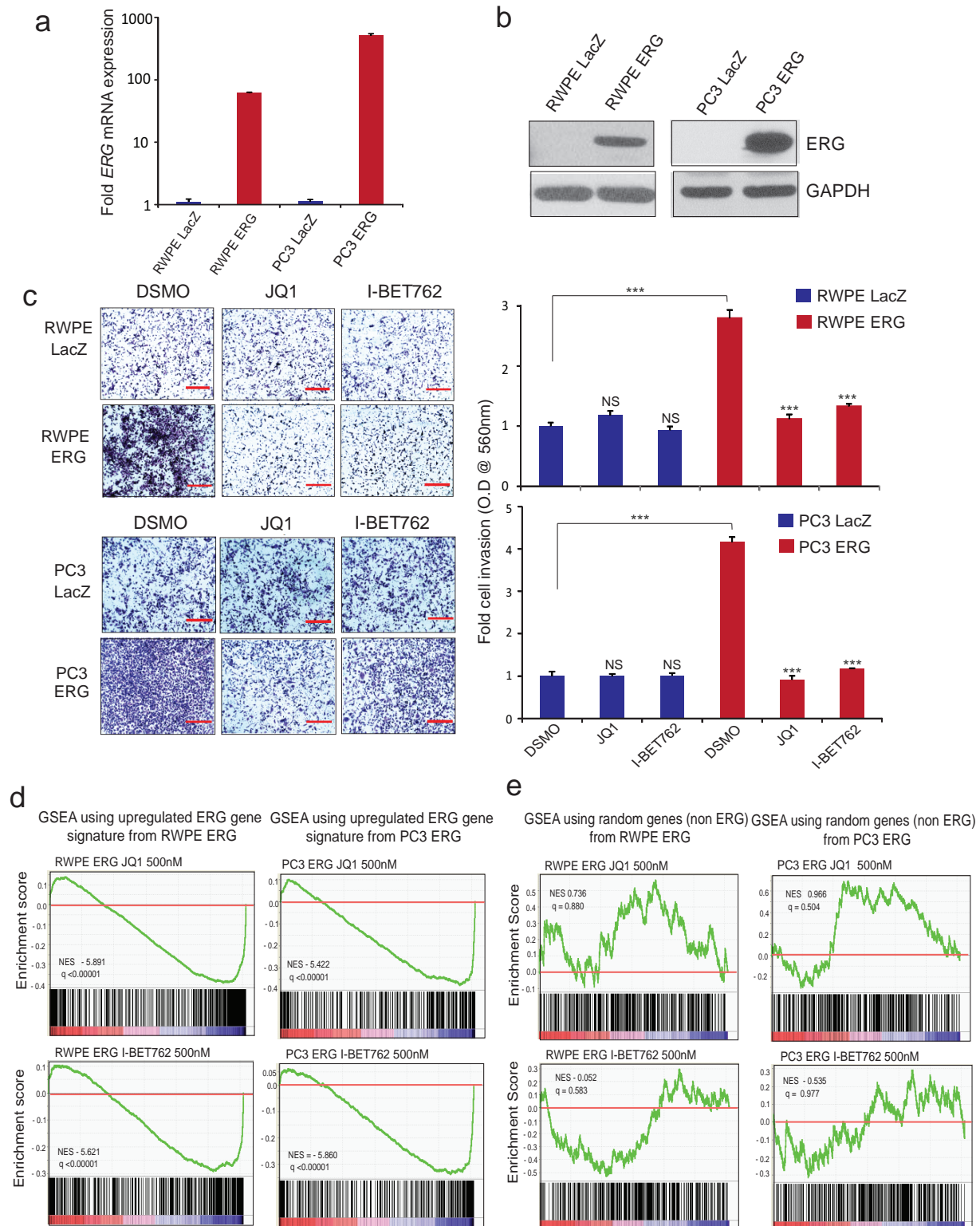
**Extended Data Figure 5 | Influence of JQ1 and anti-androgens on genome-wide recruitment of AR and their effect on DHT-induced AR target gene expression.** **a**, Two independent biological replicates of AR ChIP-seq experiments in VCaP cells show high correlation of normalized enrichment levels (see Methods) in the majority of treatment conditions.  $R^2$  values for each biological duplicate are shown. **b**, Mean enrichment (coverage) profiles are similar between biological replicates and different between treatment conditions, indicating that no adverse changes in enrichment levels are observed between the replicates. **c**, Bar graph showing total number of AR peaks for VCaP-treated cells. The genome-wide individual peaks for AR yielded the highest number of peaks for DHT (35,390) whereas vehicle control cells showed only 13,874 peaks. However, the number of peaks for AR was 23,961, 18,264 and 32,212 in the presence of JQ1, MDV3100 and bicalutamide, respectively. **d**, Heat map representation of AR binding peaks in different treatment groups. Genomic target regions are rank-ordered based on the level of AR enrichment at each androgen response elements (ARE) within -1 kb and +1 kb flanking the genomic region. **e**, Venn diagram illustrating the overlap of AR-bound genes between different treatment groups. **f**, AR-BRD4

binding on *KLK3* and *FASN* upstream regions. Genome browser representation of AR and BRD4 binding events on a putative enhancer and super-enhancer of AR-regulated *KLK3* and *FASN* gene, respectively. The  $y$ -axis denotes reads per million per base pair (r.p.m.  $\text{bp}^{-1}$ ), the  $x$ -axis denotes the genomic position with a scale bar on top right. **g**, Expression of AR target genes in the presence of JQ1, MDV3100 or bicalutamide. Heat maps for VCaP and LNCaP cells treated with DHT (10 nM), DHT plus JQ1 (0.5 μM), DHT plus MDV3100 (10 μM) and DHT plus bicalutamide (25 μM). Red arrows indicate well-characterized AR target genes. **h**, qRT-PCR analysis of AR-regulated genes in the VCaP and LNCaP treated cells. To directly compare JQ1 and MDV3100 in blocking AR signalling, cells were treated with varying concentrations of JQ1 or MDV3100 followed by DHT treatment and analysed for AR targets. The reduction in DHT-induced gene expression was observed for JQ1 even at 100–250 nM whereas MDV3100 showed a marginal reduction at 10 μM, demonstrating the higher efficacy of JQ1 in blocking AR target gene expression. Data show mean ± s.e.m. ( $n = 3$ ) from one of two independent experiments.



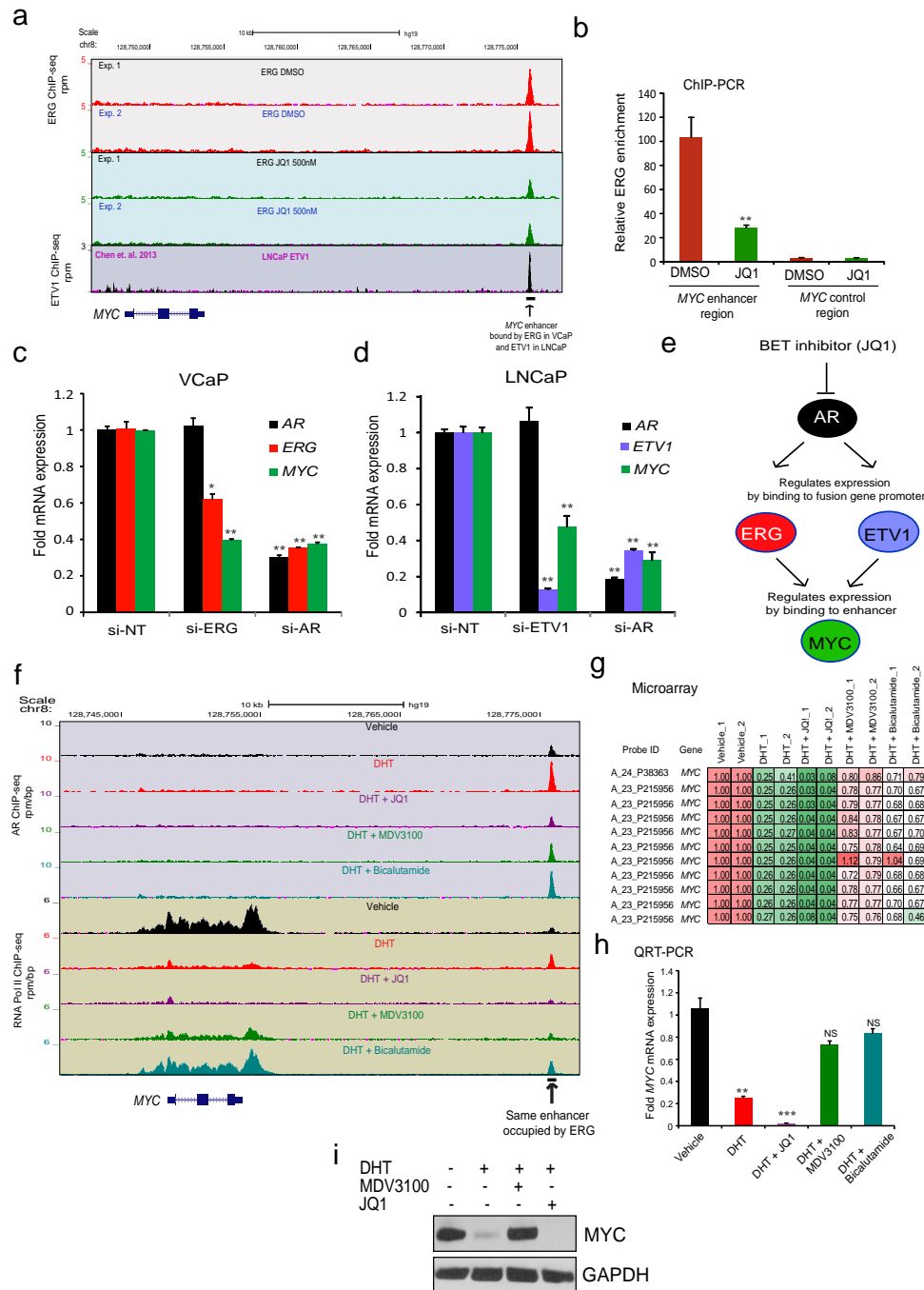
**Extended Data Figure 6 | Effect of JQ1 on the *TMPRSS2-ERG* loci and ERG-mediated transcription in VCaP cells.** **a**, Genome browser representation of RNA Pol II binding events within the *ERG* gene body. The y-axis denotes reads per million mapped reads per base pair (r.p.m. bp<sup>-1</sup>), the x-axis denotes the genomic position and the black arrow indicates the region involved in the *TMPRSS2-ERG* fusion. **b**, As in **a**, AR and BRD4 binding on the promoter of the *ERG* 5'-fusion partner *TMPRSS2* in VCaP cells. Note the reduced RNA Pol II and AR-BRD4 recruitment levels in DHT plus JQ1 tracks for the *ERG* gene body and *TMPRSS2* promoter respectively. **c**, High reproducibility of ERG ChIP-seq experiments. Biological replicates of ERG ChIP-seq experiments show very high correlation of normalized enrichment levels (see Methods) in the JQ1- and DMSO-treated conditions. **d**, Significant changes in ERG levels upon JQ1 treatment at ERG-binding sites in the proximity of gene loci. Changes in ERG enrichment levels were assessed using

DESeq2. Statistically significant differences were observed for ERG gain and ERG loss. Significant ERG gains are associated with quantitatively modest changes in enrichment level. On the other hand, significant ERG losses are associated with greater changes in enrichment levels. Individual number of peaks for each panel is shown. **e**, Genome browser representation of ERG-binding events on bona fide ERG-activating target genes. The y-axis denotes reads per million per base pair (r.p.m. bp<sup>-1</sup>), the x-axis denotes the genomic position. **f**, Genome browser representation of ERG-binding events on ERG-repressed target genes. **g**, TaqMan qRT-PCR analysis of ERG-activated genes in VCaP cells after JQ1 treatment. **h**, TaqMan QRT-PCR analysis of ERG-repressed genes in VCaP cells after JQ1 treatment. Data represent mean  $\pm$  standard deviation (s.d.) ( $n = 3$ ) from one of two independent experiments. \* $P \leq 0.05$ ; \*\* $P \leq 0.005$ ; \*\*\* $P \leq 0.0005$  by two-tailed Student's *t*-test.



**Extended Data Figure 7 | BET bromodomain inhibitors reverse ERG-mediated functions in an isogenic cell line system.** **a, b**, qRT-PCR and immunoblot showing overexpression of ERG in RWPE and PC3 prostate cell lines. Data represent mean  $\pm$  s.e.m. ( $n = 3$ ). **c**, BET inhibitors block ERG-induced RWPE and PC3 cell invasion. RWPE and PC3 cells stably expressing either LacZ or ERG were treated with DMSO ( $n = 4$ ), 500 nM JQ1 ( $n = 4$ ) or I-BET762 ( $n = 4$ ) for 24 h before plating in Matrigel-coated Boyden chambers. After 48 h cell invasion was quantified. Left, representative photomicrographs of invaded cells are shown with a 100  $\mu$ m scale bar (lower Boyden chamber stained with crystal violet). Right, bar graph shows

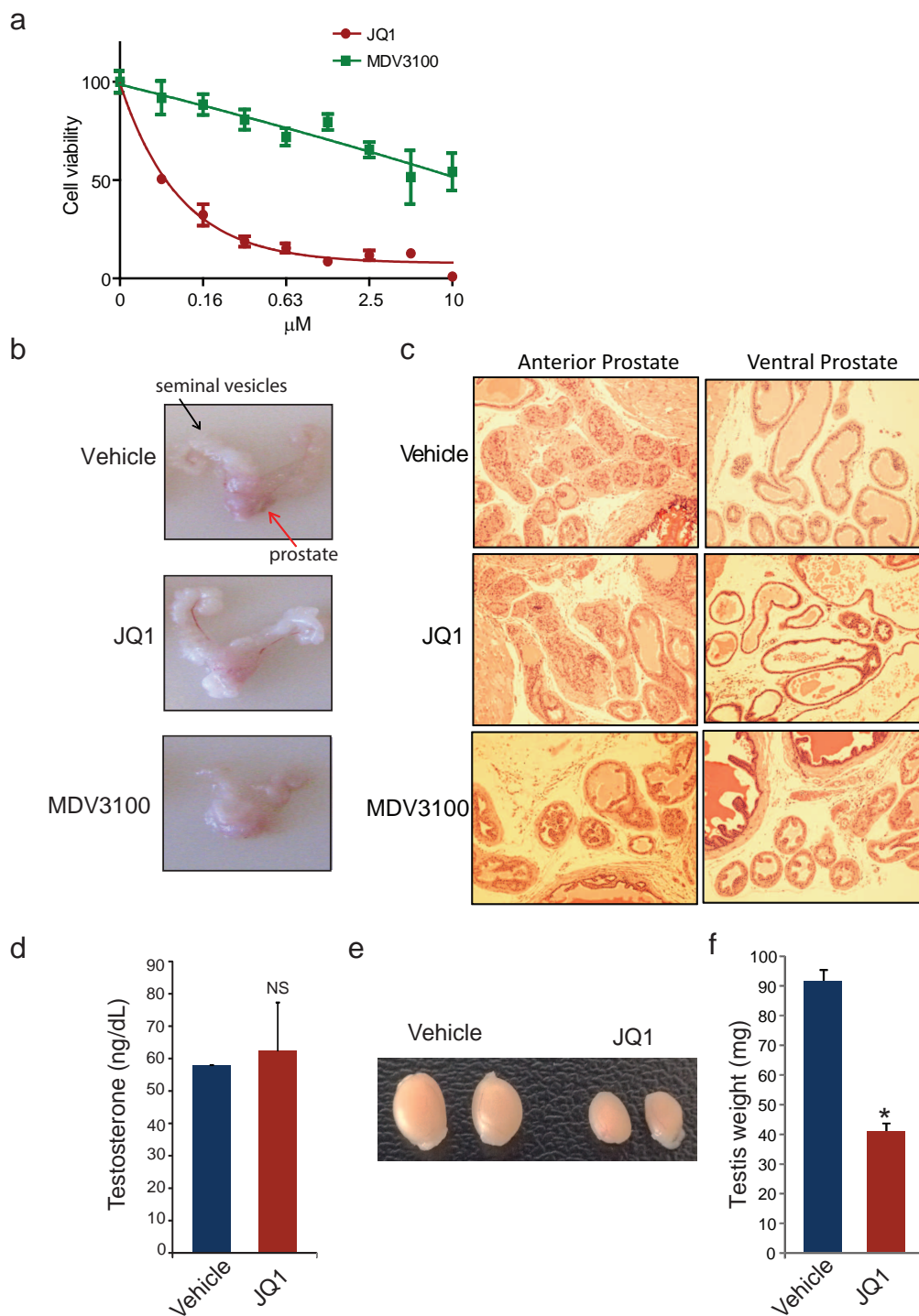
fold change in cell invasion, with DMSO-treated LacZ-expressing cells set to 1. Data represent mean  $\pm$  s.e.m. from one of three independent experiments. **d**, BET inhibitors reverse ERG-induced gene transcription. GSEA of the ERG target gene signature (see Methods) in RWPE cells overexpressing ERG (RWPE-ERG) and PC3-ERG cells treated with JQ1 or I-BET762 (500 nM) for 24 h. ERG-induced genes are repressed by JQ1 or I-BET762 treatment. **e**, GSEA using a random gene set shows no significant positive or negative enrichment by JQ1 or I-BET762 treatment in RWPE-ERG and PC3-ERG cells. NS, not significant; \*\*\* $P \leq 0.0001$  by two-tailed Student's  $t$ -test.



**Extended Data Figure 8 | JQ1 inhibits ETS (ERG/ETV1) factors that regulate MYC expression in VCaP and LNCaP cells.** **a**, Genome browser representation of ERG- and ETV1-binding events on the MYC distal enhancer<sup>42</sup>. JQ1 treatment in VCaP cells reduces ERG enrichment, as shown in two independent ERG ChIP-seq experiments. The y axis denotes reads per million per base pair (r.p.m. bp<sup>-1</sup>), the x axis denotes the genomic position. LNCaP ETV1 ChIP-seq data are based on data from ref. 23 (GEO accession code GSM1145322), and show ETV1 recruitment to the MYC distal enhancer. **b**, ChIP-PCR validation of loss of ERG recruitment after JQ1 treatment in VCaP cells. Data show mean  $\pm$  s.d. ( $n = 3$ ) from one of two independent experiments. **c**, **d**, Knock-down of AR or ETS factor reduces MYC gene expression in VCaP and LNCaP cells. qRT-PCR for AR, ETS and MYC expression in siNT, siAR or siETS transfected cells. Data show mean  $\pm$  s.d. ( $n = 3$ ) from one of two independent experiments. **e**, A cartoon illustrating the mechanism of MYC loss by JQ1 in AR-positive VCaP and LNCaP cells. **f**, Anti-androgens but not JQ1 de-repress MYC expression in prostate cancer cells. Genome browser representation of AR and RNA Pol II binding events within the MYC gene locus. The y axis denotes reads per million per base pair

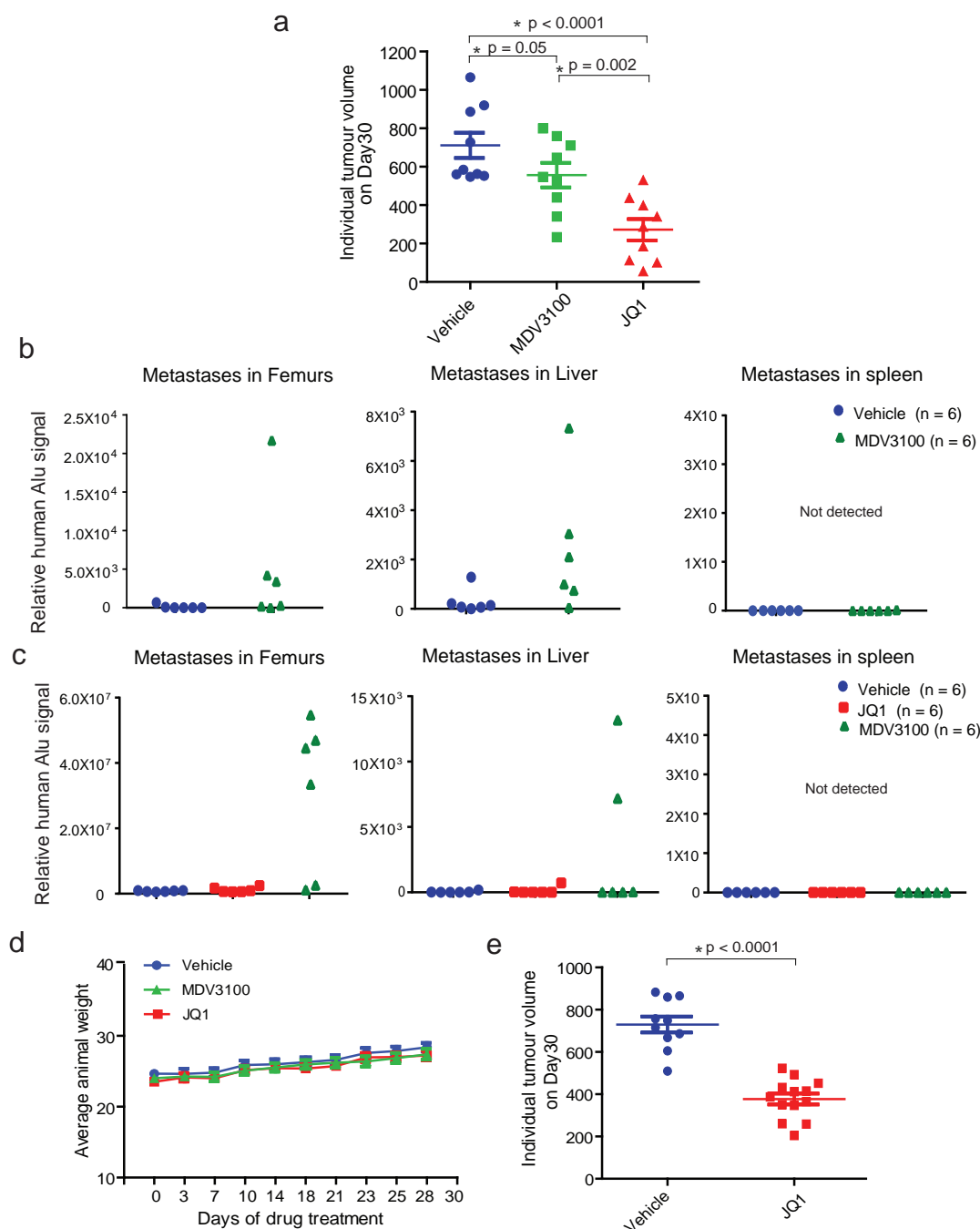
(r.p.m. bp<sup>-1</sup>), the x axis denotes the genomic position. Note the AR recruitment to the same distal enhancer that is occupied by ERG (see Extended Data Fig. 8a), indicating that there is competition between the AR and ETS factors to bind to this enhancer region to regulate MYC gene expression. **g**, Heat map showing MYC expression values from VCaP microarray gene expression data. **h**, Anti-androgen restores DHT-repressed MYC expression in VCaP cells. qRT-PCR of MYC in VCaP cells treated with vehicle, DHT (10 nM), DHT plus JQ1 (500 nM), DHT plus MDV3100 (10  $\mu$ M) or DHT plus bicalutamide (25  $\mu$ M). Inability of JQ1 to de-repress MYC in this setting could be explained by the fact that both AR and ERG are de-recruited from the MYC distal enhancer, leading to net loss of MYC expression. **i**, MDV3100 and not JQ1 restores DHT-repressed MYC protein levels in VCaP cells. Immunoblot of MYC protein in VCaP cells pre-treated with vehicle, MDV3100 (10  $\mu$ M) or JQ1 (500 nM) for 4 h followed by DHT (10 nM) for 20 h. Data show mean  $\pm$  s.d. ( $n = 3$ ) from one of two independent experiments. NS, not significant; \* $P \leq 0.01$ ; \*\* $P \leq 0.001$ ; \*\*\* $P \leq 0.0001$  by two-tailed Student's  $t$ -test.





**Extended Data Figure 9 | JQ1 does not affect normal prostate growth and testosterone levels but reduces testis size in mice.** **a**, Comparison of JQ1 and MDV3100 treatment on VCaP cell viability *in vitro*.  $N = 8$  wells of a 96-well plate per condition. VCaP cells were treated with MDV3100 or JQ1 for 8 days and assayed for viability with Cell-titerGLO. **b**, Gross images showing highly hormone-responsive seminal vesicles attached to prostate gland (red and black arrows, respectively) from male mice treated for 30 days with vehicle, JQ1 ( $50 \text{ mg kg}^{-1}$ ) or MDV3100 ( $10 \text{ mg kg}^{-1}$ ). Vehicle or JQ1-treated mice show no change in the appearance of seminal vesicles. By contrast, MDV3100-treated animals show remarkable shrinkage of seminal vesicles.

**c**, Mice treated with JQ1 do not show any adverse changes to anterior or ventral prostate morphology. The haematoxylin and eosin images show normal morphology of anterior and ventral prostate from vehicle- or JQ1-treated mice. MDV3100-treated mice show attenuated remnant glands of anterior or ventral prostate. **d**, Male mice ( $n = 3$  per group) treated with vehicle or JQ1 for 30 days exhibit similar serum testosterone levels. Data represent the mean  $\pm$  s.e.m. **e**, Gross analysis of testis from mice treated with vehicle or JQ1 for 30 days. **f**, Testis weight from vehicle control or JQ1-treated mice. Data represent the mean  $\pm$  s.e.m. from  $n = 7$  mice per group. NS, not significant;  $*P \leq 0.0001$  by two-tailed Student's *t*-test.



**Extended Data Figure 10 | In vivo effects of BET bromodomain inhibition in VCaP xenograft model.** **a**, VCaP cells were implanted subcutaneously in mice and grown until tumours reached a size of approximately 100 mm<sup>3</sup>. Xenografted mice were randomized and then received vehicle, 50 mg kg<sup>-1</sup> JQ1 or 10 mg kg<sup>-1</sup> MDV3100 5 days a week as indicated. Calliper measurements were taken twice a week. Individual tumour volumes from different treatment groups at the end of the experiments with *P* values are shown. **b**, MDV3100 treatment leads to spontaneous metastasis. Mice bearing VCaP xenografts (subcutaneously engrafted) treated with vehicle (*n* = 6) or MDV3100 (*n* = 6) were assessed for spontaneous metastasis to the femur (bone marrow) and soft tissues such as liver and spleen. Genomic DNA isolated from these sites was

analysed for metastasized cells by measuring human Alu sequence (by Alu-qPCR). MDV3100-treated mice showed spontaneous metastasis to femur and liver. Spleen did not show presence of human ALU sequences. **c**, As in **a**, for mice bearing VCaP xenografts treated with vehicle (*n* = 6), JQ1 (*n* = 6) or MDV3100 (*n* = 6). MDV3100-treated but not JQ1-treated mice showed metastasis to femur and liver. **d**, JQ1 or MDV3100 treatment does not affect animal weight. Mice from VCaP cell xenograft experiments treated with vehicle, 10 mg kg<sup>-1</sup> MDV3100 or 50 mg kg<sup>-1</sup> JQ1 were weighed at the time of calliper measurements. **e**, Individual tumour volume for vehicle- or JQ1-treated VCaP mouse xenograft (for data shown in Fig. 4c). Mean  $\pm$  s.e.m. is plotted. Statistical significance was determined by two-tailed Student's *t*-test.

# SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer

Pawel K. Mazur<sup>1,2\*</sup>, Nicolas Reynoird<sup>3\*</sup>, Purvesh Khatri<sup>4</sup>, Pascal W. T. C. Jansen<sup>5</sup>, Alex W. Wilkinson<sup>3</sup>, Shichong Liu<sup>6</sup>, Olena Barbash<sup>7</sup>, Glenn S. Van Aller<sup>7</sup>, Michael Huddleston<sup>7</sup>, Dashyant Dhanak<sup>7†</sup>, Peter J. Tummino<sup>7</sup>, Ryan G. Kruger<sup>7</sup>, Benjamin A. Garcia<sup>6</sup>, Atul J. Butte<sup>1,2</sup>, Michiel Vermeulen<sup>5†</sup>, Julien Sage<sup>1,2\*</sup> & Or Gozani<sup>3\*</sup>

**Deregulation of lysine methylation signalling has emerged as a common aetiological factor in cancer pathogenesis, with inhibitors of several histone lysine methyltransferases (KMTs) being developed as chemotherapeutics<sup>1</sup>. The largely cytoplasmic KMT SMYD3 (SET and MYND domain containing protein 3) is overexpressed in numerous human tumours<sup>2–4</sup>. However, the molecular mechanism by which SMYD3 regulates cancer pathways and its relationship to tumorigenesis *in vivo* are largely unknown. Here we show that methylation of MAP3K2 by SMYD3 increases MAP kinase signalling and promotes the formation of Ras-driven carcinomas. Using mouse models for pancreatic ductal adenocarcinoma and lung adenocarcinoma, we found that abrogating SMYD3 catalytic activity inhibits tumour development in response to oncogenic Ras. We used protein array technology to identify the MAP3K2 kinase as a target of SMYD3. In cancer cell lines, SMYD3-mediated methylation of MAP3K2 at lysine 260 potentiates activation of the Ras/Raf/MEK/ERK signalling module and SMYD3 depletion synergizes with a MEK inhibitor to block Ras-driven tumorigenesis. Finally, the PP2A phosphatase complex, a key negative regulator of the MAP kinase pathway, binds to MAP3K2 and this interaction is blocked by methylation. Together, our results elucidate a new role for lysine methylation in integrating cytoplasmic kinase-signalling cascades and establish a pivotal role for SMYD3 in the regulation of oncogenic Ras signalling.**

The Ras family of oncogenes is activated in a large fraction of human cancers<sup>5</sup>. To explore possible connections between KMTs and human cancers driven by activation of Ras, we surveyed the level of expression for 54 known and candidate human KMT genes in pancreas ductal adenocarcinoma (PDAC), a cancer nearly universally initiated by oncogenic Ras mutations. We found that five KMT-encoding genes (*SMYD3*, *MLL5* (also known as *KMT2E*), *EZH2*, *SETD5* and *WHSC1L1*) were consistently upregulated in human data sets (Extended Data Fig. 1a–c). *SMYD3*, which showed the most significant correlation with PDAC in our meta-analysis, is reported to be overexpressed in several cancers with elevated Ras activity<sup>2,3</sup>, and *SMYD3* protein expression increases with PDAC development (Extended Data Fig. 2a, b). On the basis of these results, we postulated that *SMYD3* might have a role in Ras-driven tumorigenesis.

Little is known regarding *SMYD3* cancer-related function *in vivo*. We previously generated *Smyd3* mutant mice, which develop normally, and are viable and fertile (Extended Data Fig. 2d; data not shown; ref. 6). To investigate the role of *SMYD3* in Ras-driven cancers, *Smyd3* mutant mice were crossed with mice harbouring a *loxP-Stop-loxP Kras<sup>G12D</sup>* knock-in allele (*Kras<sup>+/LSL-G12D</sup>*), which allows for the controlled induction of oncogenic K-Ras and the initiation of tumours in distinct organs, including the lung and the pancreas<sup>7,8</sup>. PDAC is thought to arise from the transdifferentiation of acinar cells into duct-like cells upon activation

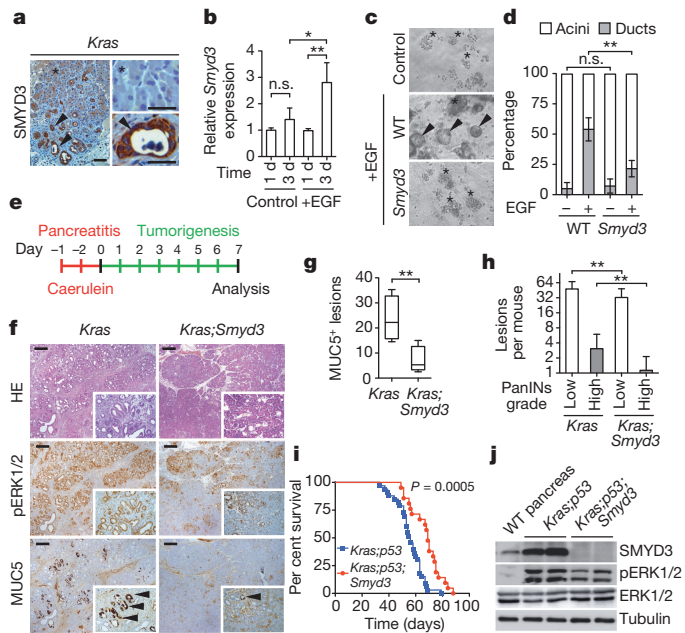
of Ras signalling<sup>9,10</sup>. We observed induction of *SMYD3* expression during this process in mice with pancreas-specific activation of K-Ras (Fig. 1a) and in an *ex vivo* acinar-to-ductal metaplasia (ADM) assay<sup>11</sup> (Fig. 1b). In this assay, *SMYD3* was required for efficient duct formation from acinar cells (Fig. 1c, d). *In vivo*, ADM and PDAC initiation are triggered in young *Kras* mutant mice by inducing severe acute pancreatitis via repeated injections of caerulein<sup>12</sup> (Fig. 1e). In this system, *Smyd3* deletion reduced the appearance of pancreatic intra-epithelial neoplasia (PanIN) brought on by *Kras* activation, as determined by histopathological analysis and decreased signal for both phosphorylated ERK1/2 (pERK1/2, a downstream marker of Ras activity) and MUC5 (a marker of PanINs) (Fig. 1f, g). In the absence of a pancreatitis trigger, PanIN lesions develop by 6 months in *p48<sup>+/Cre</sup>;Kras<sup>+/LSL-G12D</sup>* mice<sup>8</sup>, a process that was attenuated by *Smyd3* loss (Fig. 1h; Extended Data Fig. 3a). Next, to study PDAC growth and to perform survival studies, we used the *p48<sup>+/Cre</sup>;Kras<sup>+/LSL-G12D</sup>;p53<sup>lox/lox</sup>* (*Kras;p53*) mutant model (*p48* and *p53* are also known as *Ptf1a* and *Trp53*, respectively), which is characterized by rapid PanIN-to-PDAC progression and malignant transformation with 100% penetrance in a relatively short latency (~50–60 days)<sup>13</sup>. At autopsy, the pancreatic tissue from *Kras;p53* mutant mice was entirely occupied by transformed cells, whereas areas of normal pancreatic tissue remained in *Kras;p53;Smyd3* mutant mice (Extended Data Fig. 3b, c). Furthermore, loss of *Smyd3* extended the lifespan of the animals (Fig. 1i) and resulted in reduced levels of the PDAC biomarker pERK1/2 in biopsy samples (Fig. 1j; Extended Data Fig. 3b). Notably, K-Ras expression was not affected by *SMYD3* deletion (Extended Data Fig. 3f). Based on these data, we conclude that *SMYD3* is required for efficient initiation of pancreatic cancer by oncogenic K-Ras.

Oncogenic activation of the Ras pathway is a frequent event in lung adenocarcinoma, a cancer that also shows high *SMYD3* expression (Extended Data Figs 1d and 2c). Intratracheal injection of an adenovirus expressing the Cre recombinase (Ad-Cre) in adult *Kras<sup>+/LSL-G12D</sup>* mice led to the development of atypical adenomatous hyperplasia (AAH) and adenomas in the lungs within 12 weeks<sup>7</sup>, irrespective of *Smyd3* status (Fig. 2a, b; Extended Data Fig. 3d). In contrast, at 16 and 20 weeks or more after Ad-Cre infection, mice lacking *Smyd3* showed significantly smaller and less advanced tumours than control mice (Fig. 2a, c, d; Extended Data Fig. 3d; data not shown). Specifically, quantification of tumour grade indicated that *Smyd3* loss impeded the critical transition from adenoma to adenocarcinoma (Fig. 2c), which was also observable at the whole-organ level (Fig. 2d). Moreover, the lifespan of *Kras<sup>G12D</sup>*-expressing mice was 20% longer if they were mutant for *Smyd3* (Fig. 2e). Progression of lung cancer to carcinoma correlates with amplification of Ras/MEK/ERK signalling<sup>14,15</sup>. *Smyd3* deletion resulted in lower detection of pERK1/2 relative to control tumours without an overall change

<sup>1</sup>Department of Pediatrics, Stanford University School of Medicine, California 94305, USA. <sup>2</sup>Department of Genetics, Stanford University School of Medicine, California 94305, USA. <sup>3</sup>Department of Biology, Stanford University, California 94305, USA. <sup>4</sup>Institute for Immunity, Transplantation and Infection, and Department of Medicine, Stanford University School of Medicine, California 94305, USA.

<sup>5</sup>Department of Molecular Cancer Research and Department of Medical Oncology, University Medical Center Utrecht, 3508 AB Utrecht, The Netherlands. <sup>6</sup>Epigenetics Program and Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. <sup>7</sup>Cancer Epigenetics DPU, Oncology R&D, GlaxoSmithKline, Collegeville, Pennsylvania 19426 USA. <sup>†</sup>Present addresses: Janssen Research and Development, 1400 McKean Road, Spring House, Pennsylvania 19477, USA (D.D.); Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University, 6525GA Nijmegen, The Netherlands (M.V.).

\*These authors contributed equally to this work.



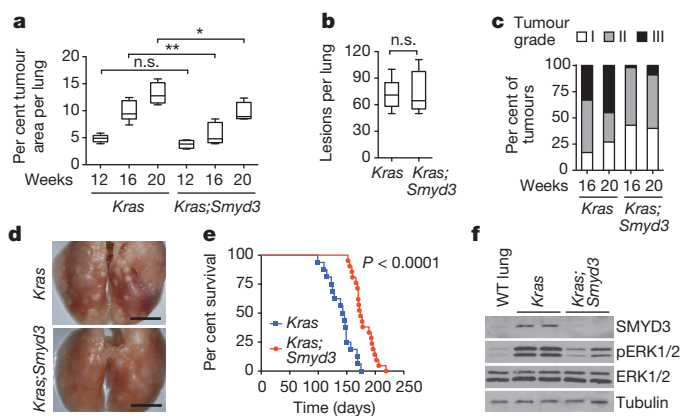
**Figure 1 | SMYD3 loss inhibits Ras-driven pancreatic tumorigenesis.**

**a**, Representative immunohistochemistry (IHC) images showing SMYD3 expression in cells undergoing acinar-to-ductal metaplasia (ADM, arrowheads) but not in acini (asterisk) in  $p48^{Cre/+};Kras^{+}/LSL-G12D$  (*Kras*) mice. **b**, *Smyd3* expression increases during ADM formation. Quantitative PCR with reverse transcription (qRT-PCR) analysis of *Smyd3* expression at the indicated times from control- and EGF-induced ADM *ex vivo* samples (four independent biological replicates). **c**, SMYD3 depletion inhibits ADM. Wild-type (WT,  $p48^{Cre/+}$ ) acinar clusters (asterisk) undergo ADM and form ducts (arrowhead) *ex vivo*, whereas *Smyd3* mutant acini explants inefficiently form ducts. **d**, Quantification of acinar and ductal clusters on day 3 of culture as in **c** (four independent biological replicates with three technical replicates each). **e**, Schematic of the caerulein pancreatitis-induced tumorigenesis protocol<sup>12</sup>. **f**, Representative haematoxylin and eosin (HE) staining and IHC for pERK1/2, a marker of Ras activity, and MUC5, a marker of PanIN lesions (arrowheads). **g**, Quantification of MUC5-positive lesions in caerulein-treated pancreata from *Kras* ( $n = 6$ ) and *Kras;Smyd3* ( $n = 6$ ) mutant mice. **h**, Quantification of spontaneous PanIN lesions formed in 6-month-old *Kras* ( $n = 8$ ) and *Kras;Smyd3* ( $n = 8$ ) mutant mice. The grade of lesions is indicated. **i**, Kaplan-Meier survival of *Kras;p53* mutant mice ( $p48^{Cre/+};Kras^{+}/LSL-G12D/+;p53^{lox/lox}$ ,  $n = 33$ , median survival = 56 days) and *Kras;p53;Smyd3* mutant mice ( $n = 21$ , median survival = 68.5 days) animals.  $P = 0.0005$  by log-rank test for significance. **j**, Immunoblots with the indicated antibodies of *Kras;p53* and *Kras;p53;Smyd3* mutant pancreatic tumour lysates. Loss of SMYD3 was also confirmed by immunostaining of pancreatic sections (Extended Data Fig. 4d). All scale bars, 50  $\mu$ m. \* $P$  value < 0.05; \*\* $P$  value < 0.01; n.s., not significant (two-tailed unpaired Student's *t*-test). Data are represented as mean  $\pm$  s.e.m.

in total levels of Ras (Fig. 2f; Extended Data Figs 2a and 3f). Together, these observations indicate that SMYD3 promotes Ras-driven cancer development and progression *in vivo*.

Depletion of SMYD3 by RNA-mediated interference (RNAi) using a short hairpin RNA (shRNA) strategy in LKR10 mouse cells (a LAC-derived cell line<sup>16</sup>), A549 (a human LAC cell line) and CFPac1 (a human PDAC cell line) reduced the proliferation rates for all three cell types, and inhibited their ability to grow in anchorage-independent condition (Extended Data Fig. 4a–c). Furthermore, knockdown of SMYD3 in CFPac1 cells inhibited tumour growth in mouse xenograft experiments (Extended Data Fig. 4d–f). Thus, SMYD3 acts to maintain a number of tumorigenic characteristics in mouse and human cancer cell lines driven by oncogenic Ras.

Next, wild-type SMYD3, a catalytically inactive form (SMYD3(F183A))<sup>6</sup>, or vector control were co-expressed with the Cre recombinase in the lungs of *Kras;Smyd3* mutant mice by lentiviral transduction (Extended Data Fig. 5a). Complementation of wild-type SMYD3 into the lungs of



**Figure 2 | SMYD3 loss inhibits the development of Ras-driven lung adenocarcinoma.**

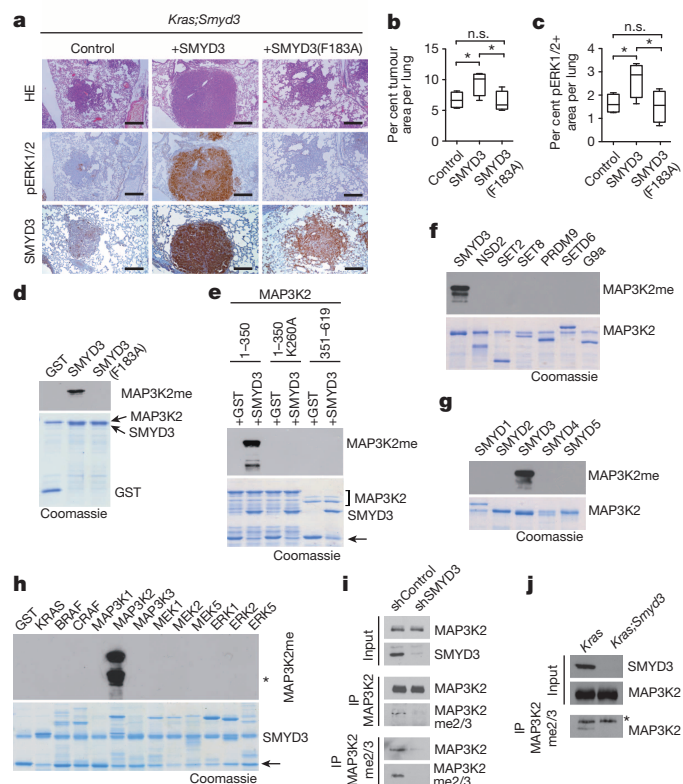
**a**, Quantification of tumour area per lung at the indicated time points after tumour induction ( $n = 6$  for each time point and genotype). **b**, Total number of tumour lesions at 12 weeks post tumour induction ( $n = 6$  for each genotype). Data are represented as mean  $\pm$  s.e.m. \* $P$  value < 0.05; \*\* $P$  value < 0.01; n.s., not significant (two-tailed unpaired Student's *t*-test). **c**, Quantification of tumour grade at 16 and 20 weeks ( $n = 6$  for each time point and genotype). **d**, Representative lung images at the endpoint of survival studies. Scale bars, 1 cm. **e**, Survival analysis of *Kras* ( $n = 16$ , median survival = 144.5 days) and *Kras;Smyd3* ( $n = 21$ ; median survival = 174 days) mutant mice, time post infection.  $P < 0.0001$  by log-rank test for significance. **f**, Immunoblots of lung tumour lysates dissected from *Kras* and *Kras;Smyd3* mutant mice (two independent biological replicates for each genotype) with the indicated antibodies. Loss of SMYD3 was also confirmed by immunostaining on lung sections (Extended Data Fig. 3f).

*Kras;Smyd3* mutant mice resulted in a higher tumour burden and pERK1/2 signal relative to the control Cre-alone infection and expression of mutant SMYD3(F183A) (Fig. 3a–c). Reconstitution experiments also demonstrated that SMYD3 catalytic activity is required for pancreatic ADM (Extended Data Fig. 5b, c). We previously reported that SMYD3 methylates histone H4 at lysine 5 (H4K5) and not at other lysines on histones<sup>6</sup>. However, in LAC and PDAC cells and tumours, virtually all SMYD3 is present in the cytoplasm (Fig. 1a and Extended Data Figs 2a, c, 3e and 6b), indicating that the cancer-relevant substrate in these cell types is unlikely to be nuclear H4K5 but rather a cytoplasmic protein.

In a biochemical screen for SMYD3 targets on a protein array platform containing more than 9,000 potential substrates<sup>17</sup>, the only candidate to be methylated by SMYD3 in three independent experiments was the MAP kinase pathway component MAP3K2 (Extended Data Fig. 6a). As shown in Fig. 3d, recombinant wild-type SMYD3, but not the catalytically dead SMYD3(F183A) mutant, methylated recombinant MAP3K2 *in vitro*. Using a mutagenesis approach, we identified lysine 260 of MAP3K2 as the single site of methylation catalysed by SMYD3 (Fig. 3e; data not shown). The immediate sequence surrounding K260 of MAP3K2 and K5 of H4 is identical (GKGG), although the catalytic efficiency ( $k_{cat}/K_M$ ) of SMYD3 for MAP3K2 is nearly two orders of magnitude greater than it is for H4 (Extended Data Fig. 6c). We also did not detect any methylation of H3 at lysine 4 by SMYD3, a previously reported activity<sup>4</sup> (Extended Data Fig. 6d, e). In addition, SMYD3 was the only KMT of the eleven we tested that could methylate MAP3K2 (Fig. 3f, g; Extended Data Fig. 6f). Furthermore, whereas SMYD3 methylated MAP3K2, it had no detectable activity on a dozen other members of the MAP kinase signalling cascade (Fig. 3h). Finally, *in vitro* methylation assays on MAP3K2 peptides spanning K260 (amino acids 249 to 273) with K260 either unmethylated, mono-, di- or tri-methylated showed that SMYD3 can use all lower states of methylation as substrates to generate the fully saturated trimethyl state at K260 (Extended Data Fig. 6g, h). Thus, SMYD3 mono-, di- and tri-methylates MAP3K2 at lysine 260 (MAP3K2-K260me) with high specificity *in vitro*.

To investigate MAP3K2 methylation in cells, we raised methyl-specific antibodies against the different states of methylation at K260 (Extended





**Figure 3 | SMYD3 methylates MAP3K2 in cancer cells.** **a**, Analysis of lung cancer development in *Kras;Smyd3* mutant mice following infection with a lentivirus expressing Cre only or simultaneously Cre and WT SMYD3 or inactive SMYD3(F183A). Histological analysis (HE staining) and IHC for pERK1/2 were performed 24 weeks post lentiviral infection. IHC confirms lentiviral-mediated expression of SMYD3. Scale bars, 50  $\mu$ m. **b**, Quantification of total tumour area per lung and pERK1/2 positive area per lung, respectively ( $n = 4$  for each experimental group). Data are represented as mean  $\pm$  s.e.m. \* $P$  value  $< 0.05$  (two-tailed unpaired Student's  $t$ -test). **d**, SMYD3 directly methylates MAP3K2. *In vitro* methylation assay on full-length recombinant MAP3K2 with recombinant wild-type SMYD3, catalytic-dead SMYD3(F183A), or glutathione-S-transferase (GST) control. Top panel, autoradiogram of methylation assay. Bottom panel, Coomassie stain of proteins in the reaction. **e**, SMYD3 methylates MAP3K2 at K260. *In vitro* methylation assay as in **d** with the indicated proteins on MAP3K2 amino acids 1–350, MAP3K2 amino acids 1–350 with a K260A substitution, and MAP3K2 amino acids 351–619. Arrow indicates GST, which is a stable breakdown product of recombinant proteins. **f**, **g**, MAP3K2 is a specific substrate of SMYD3. *In vitro* methylation assays as in **d** on MAP3K2 using the indicated KMTs (positive controls for the known KMTs shown in Extended Data Fig. 6f). **h**, SMYD3 specifically methylates MAP3K2. *In vitro* SMYD3 methylation assay as in **d** on the indicated MAP kinase pathway proteins. Asterisk indicates MAP3K2 breakdown product. Arrow indicates GST, which is a stable breakdown product for many of the screened substrates. **i**, Immunoblots with the indicated antibodies from input (cytoplasmic extract) or the indicated IPs (immunoprecipitations) from LKR10 cells stably expressing control or *Smyd3* shRNA. **j**, Immunoblots with the indicated antibodies and samples as in **i** of lung tumour biopsy lysates isolated from *Kras* and *Kras;Smyd3* mutant mice. Asterisks represent detection of IgG. For experiments **e–k** representative data based on three or more independent biological replica are shown.

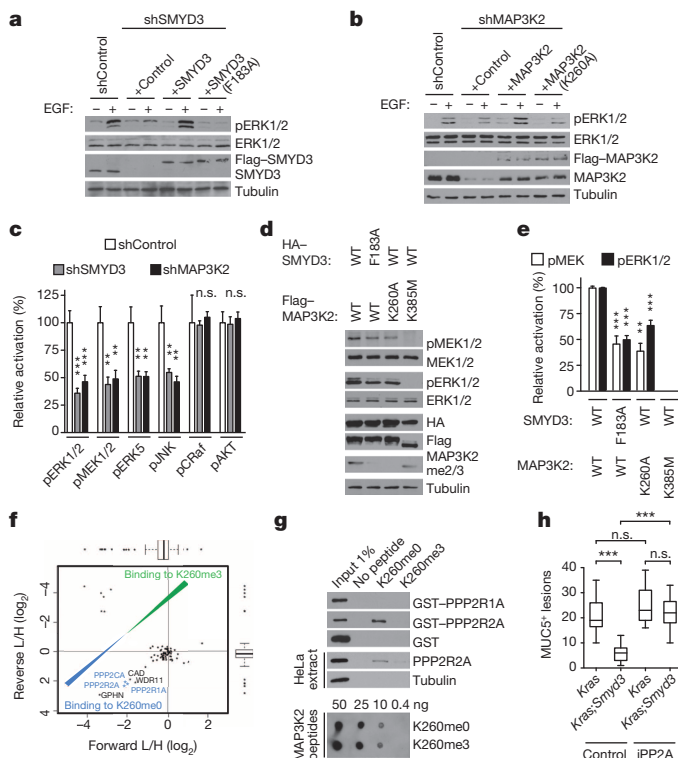
Data Fig. 6i). In co-transfection experiments in human 293T cells, overexpressed MAP3K2 was methylated at K260 upon SMYD3 overexpression (Extended Data Fig. 6j). Endogenous methylation at MAP3K2-K260 was observed in LKR10 cells and RNAi-mediated depletion of SMYD3 in these cells resulted in loss of this signal (Fig. 3i). Finally, the MAP3K2-K260me2/3 signal was significantly reduced in tumour tissue microdissected from *Kras* versus *Kras;Smyd3* mutant mice (Fig. 3j). Thus,

SMYD3 is required for maintenance of physiological levels of MAP3K2-K260 methylation in cancer tissue and cells.

The cytoplasmic kinase MAP3K2 is activated in response to a variety of stress and mitogenic stimuli, including epidermal growth factor (EGF), and relays signals to downstream MAP kinase components such as ERK1/2<sup>18</sup> and ERK5<sup>19</sup>. Given that the pERK1/2 and pERK5 signals are reduced in LAC and PDAC samples from *Smyd3* mutant mice (Figs 1 and 2; Extended Data Fig. 3), we reasoned that SMYD3-mediated methylation of MAP3K2 may regulate signalling within the Ras/ERK pathway. We therefore examined the relationship between SMYD3 and EGF-stimulated ERK1/2 activation. First, endogenous SMYD3 was depleted in LKR10 cells by shRNA targeting the 3' untranslated region of *Smyd3* and then we reconstituted the depleted cells with RNAi-resistant wild-type SMYD3 or catalytically inactive SMYD3(F183A). In control cells, EGF treatment triggered ERK1/2 phosphorylation, and this response was greatly reduced by SMYD3 depletion (Fig. 4a). Complementation with wild-type SMYD3 re-established the EGF-mediated ERK1/2 phosphorylation response, whereas complementation with SMYD3(F183A) failed to do so (Fig. 4a). SMYD3 was also required when serum was used to activate ERK1/2 activation in LKR10 cells (Extended Data Fig. 7b) and when EGF was used to activate ERK1/2 in human lung and pancreatic cancer cells (Extended Data Fig. 7c, d). Finally, we established a MAP3K2 complementation system to investigate the role of K260 of MAP3K2. In these experiments, wild-type MAP3K2, but not the SMYD3-resistant MAP3K2(K260A) mutant, reconstituted the EGF-mediated ERK1/2 phosphorylation response (Fig. 4b).

To characterize how the SMYD3-MAP3K2 axis impinges on the overall MAP kinase network, the level of EGF-induced activation for several kinases was determined in control, SMYD3 knockdown and MAP3K2 knockdown LKR10, A549 and CFPac1 cells (Fig. 4c; Extended Data Fig. 7a–d). SMYD3 and MAP3K2 were both required for full activation of ERK5, ERK1/2, MEK1/2 and JNK (known downstream targets of MAP3K2; refs 18–22), but dispensable for activation of AKT and Craf (RAF-1) (p38 was not activated in the cell lines tested; Extended Data Fig. 7a–d). Notably, MEK1/2 activation was impaired in SMYD3 and MAP3K2 knockdown cell lines. As the canonical ERK1/2 activation pathway consists of Raf-MEK1/2-ERK1/2 and both SMYD3 and MAP3K2 were required for MEK1/2 and ERK1/2 activation but not Craf, we postulated that SMYD3 methylation of MAP3K2 directly influences MEK1/2 phosphorylation. Indeed, pMEK1/2 signal increased in response to overexpression of wild-type MAP3K2 and a SMYD3-resistant MAP3K2(K260A) mutant, but not a catalytically dead MAP3K2(K385M) mutant<sup>23</sup> (Fig. 4d, e; Extended Data Fig. 7e). SMYD3 expression alone had no effect on MEK1/2 activation (Extended Data Fig. 7e), but resulted in increased MEK1/2 phosphorylation when co-expressed with wild-type MAP3K2 relative to co-expression with MAP3K2(K260A) (Fig. 4d, e). Notably, MAP3K2 phosphorylates MEK1 *in vitro* kinase assays<sup>21</sup>, and this activity was unchanged if SMYD3-methylated MAP3K2 or MAP3K2(K260A) were used in the kinase assays rather than wild-type, unmethylated MAP3K2 (Extended Data Fig. 7f). These results indicate that the molecular mode of action linking MAP3K2 methylation to MEK1/2 activation is not due to changes in the intrinsic kinase activity of MAP3K2 but rather another mechanism.

Given the role of SMYD3-MAP3K2 in activating MEK1/2, we tested whether SMYD3 depletion augments the effects of the MEK1/2 inhibitor trametinib, which is currently being investigated to treat NSCLC and PDAC (<http://clinicaltrials.gov/>). Administration of *Kras* and *Kras;Smyd3* mutant mice with a normal dose of trametinib blocked tumorigenesis in both strains, although phosphorylation of ERK1/2 was still lower in mice depleted of SMYD3 (Extended Data Fig. 8). Notably, a low-dose trametinib regimen, which only partially inhibited pERK1/2 levels and the formation of neoplastic lesions in *Kras* mutant mice, was sufficient to block tumorigenesis and ERK1/2 activation in *Smyd3* knockouts (Extended Data Fig. 8). Trametinib was also more potent in cancer cell lines when coupled with SMYD3 depletion (Extended Data Fig. 9a). These data indicate that SMYD3 may act in concert with MEK1/2 signalling



in tumorigenesis. Indeed, overexpression of activated MEK1(S218D, S222D) (MEK1-DD)<sup>24</sup> rescued the effects of SMYD3 depletion in lung cancer cells (Extended Data Fig. 9b)

Because the intrinsic kinase activity of MAP3K2 is not directly altered by methylation, we postulated that this modification event was involved in modulating a key protein–protein interaction. To identify candidate methyl-sensitive binding partners of MAP3K2, a SILAC (stable isotope labelling by amino acids in cell culture)-based quantitative proteomic screen was performed with cytoplasmic extracts to isolate proteins that bound differentially to MAP3K2-K260me0 peptides versus MAP3K2-K260me3 peptides. Although this analysis did not show enrichment of any K260me-binding proteins, it did reveal six candidates that bind to the unmethylated peptide but are blocked by K260 trimethylation. Of these six proteins, three are members of the PP2A serine/threonine phosphatase complex (Fig. 4f). In our proteomics experiment we identified the catalytic PPP2CA protein, the structural PPP2R1A protein and the regulatory PPP2R2A protein, three classes of subunits that comprise the typical heterotrimeric PP2A complex. This complex is a major cellular phosphatase that inactivates key members of the MAP kinase signalling cascade (reviewed in refs 25, 26). We found that the interaction between the PP2A complex and MAP3K2 is direct, as recombinant PPP2R2A, but not PPP2R1A, bound to MAP3K2-K260me0 peptides and not to MAP3K2-K260me3 peptides in *in vitro* peptide pull-down experiments (Fig. 4g; top panel). Moreover, PPP2R2A was specifically recovered from cytoplasmic extracts using MAP3K2-K260me0 peptides but not by MAP3K2-K260me3 peptides (Fig. 4g; middle panel). Thus, amino acids 249–273 of MAP3K2 are sufficient for binding directly to the PP2A complex, via PPP2R2A, and methylation at K260 inhibits this interaction.

We next tested the ability of the PP2A inhibitor cantharidin<sup>27</sup> to ‘phenocopy’ SMYD3 function (Fig. 4h; Extended Data Fig. 10a–d). Cantharidin treatment had no gross effect on tumour formation in *p48<sup>+/Cre</sup> Kras* mutant mice relative to vehicle treatment; in contrast, administration of cantharidin to *p48<sup>+/Cre</sup> Kras;Smyd3* mutant mice restored tumour formation to the level seen with wild-type SMYD3 in the *p48<sup>+/Cre</sup> Kras* mutant mice. These data suggest an *in vivo* functional connection between the Ras pathway, SMYD3 and PP2A.

We have identified SMYD3-catalysed methylation of MAP3K2 as a key event regulating Ras signalling in cancer cells. Although MAP3K2

**Figure 4 | SMYD3 methylation of MAP3K2 activates MAP kinase signalling pathways and repels PP2A.** **a**, SMYD3 catalytic activity is required for ERK1/2 activation in LAC cells. Immunoblots with the indicated antibodies of LAC cells lysates depleted for SMYD3 and reconstituted with either active or inactive Flag-SMYD3 as indicated. Stimulation, EGF treatment for 15 min at  $25 \text{ ng } \mu\text{l}^{-1}$ . **b**, MAP3K2 methylation is required for ERK1/2 activation in LAC cells. Immunoblot with the indicated antibodies of LAC cells lysates treated as in **a** depleted for MAP3K2 and reconstituted with either wild-type or K260A mutant Flag-MAP3K2. **c**, SMYD3 and MAP3K2 regulate multiple overlapping MAP kinase pathway proteins. Quantification of the indicated activated kinase signals in shSMYD3 and shMAP3K2 relative to shControl cell lysates based on three independent biological replica treated as in Extended Data Fig. 7a. **d**, SMYD3 catalytic activity promotes MAP3K2-induced phosphorylation of MEK1/2. Immunoblots with the indicated antibodies of 293T cells lysates transfected with Flag-MAP3K2 and haemagglutinin (HA)-SMYD3 wild-type and derivatives as indicated. MAP3K2(K385) is a kinase-dead mutant. **e**, Quantification of pMEK and pERK1/2 signals in 293T cells transfected with the indicated MAP3K2 and SMYD3 constructs and treated as in **d**. Data were generated from three independent biological replica. **f**, SILAC-based quantitative proteomic analysis of proteins that bind to MAP3K2-K260me0 and MAP3K2-K260me3 peptides. Data represent two independent experiments (forward and reverse direction). Proteins are plotted by their SILAC ratios in the forward ( $x$  axis) and reverse ( $y$  axis) SILAC experiments. Specific interactors of K260me0 reside in the lower left quadrant. The three PP2A complex components are highlighted in blue. L/H, light over heavy fraction ratio. **g**, PP2R2A directly binds to MAP3K2 peptides encompassing amino acids 249–273 and this interaction is inhibited by K260 methylation. Immunoblots of peptide pull-downs as indicated with either recombinant proteins (top panel) or HeLa cytoplasmic extracts (middle panel) (based on two replica). The dot blot in the bottom panel shows equivalent amounts of peptides used for the experiments. **h**, Quantification of MUC5+ lesions in caerulein-treated pancreata from *Kras* ( $n = 5$ , each treatment) and *Kras;Smyd3* ( $n = 5$ , each treatment) mutant mice treated with the PP2A inhibitor cantharidin (iPP2A) ( $0.15 \text{ mg kg}^{-1}$  twice a day, intraperitoneally) or vehicle control (see Extended Data Fig. 10). \*\* $P$  value  $< 0.01$ ; \*\*\* $P$  value  $< 0.001$ ; n.s., not significant (two-tailed unpaired Student’s  $t$ -test). Data are represented as mean  $\pm$  s.e.m.

was shown to phosphorylate MEK1/2 directly, a complete understanding of how MAP3K2 functions in Ras signalling remains to be determined. Our data suggest a model in which increased SMYD3 activity generates a population of methylated MAP3K2, which—via mechanisms such as blocking the association of the PP2A phosphatase with components of the MAP kinase network—intensifies the output of this pathway in response to oncogenic Ras (Extended Data Fig. 10e). In addition, the cytoplasmic localization of the SMYD3-MAP3K2 dynamic suggests a paradigm for how signalling through lysine methylation and phosphorylation can be integrated to regulate key signal transduction cascades. A clinical implication of this work is the identification of SMYD3 as a candidate therapeutic target for pharmacologic intervention to treat pancreatic and lung cancers, as well as potentially other Ras-driven tumours. The complete loss of SMYD3 function has no visible phenotype in mice, suggesting that SMYD3 inhibitors would have minimal collateral toxicity as chemotherapeutics. Thus, one could envision a therapeutic strategy comprising inhibitors of Raf or MEK that are currently used in the clinic with a SMYD3 inhibitory agent, which could mitigate potential drug toxicity by lowering the overall dose needed for each medicine and combat the development of resistance. Together, our findings reveal a new function for lysine methylation signalling in the cytoplasm in the regulation of cancer pathways.

## METHODS SUMMARY

Animal studies were performed according to practices prescribed by the NIH at Stanford’s Research Animal Facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care. The *Smyd3* mutant mice used in lung and pancreatic cancer experiments were initially obtained from the KOMP Repository. SILAC was performed using HeLa cells extracts grown in either normal amino acids culture condition (‘light’) or using modified amino acids culture condition (‘heavy’). Full methods and data analysis are described in detail in Methods.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 25 September 2013; accepted 11 April 2014.**

**Published online 21 May; corrected online 11 June 2014 (see full-text HTML version for details).**

1. Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488 (2013).
2. Watanabe, T. *et al.* Differential gene expression signatures between colorectal cancers with and without KRAS mutations: crosstalk between the KRAS pathway and other signalling pathways. *Eur. J. Cancer* **47**, 1946–1954 (2011).
3. Gaedcke, J. *et al.* Mutated KRAS results in overexpression of DUSP4, a MAP-kinase phosphatase, and SMYD3, a histone methyltransferase, in rectal carcinomas. *Genes Chromosom. Cancer* **49**, 1024–1034 (2010).
4. Hamamoto, R. *et al.* SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nature Cell Biol.* **6**, 731–740 (2004).
5. Pylayeva-Gupta, Y., Grabocka, E. & Bar-Sagi, D. RAS oncogenes: weaving a tumorigenic web. *Nature Rev. Cancer* **11**, 761–774 (2011).
6. Van Aller, G. S. *et al.* Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics* **7**, 340–343 (2012).
7. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
8. Hingorani, S. R. *et al.* Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* **4**, 437–450 (2003).
9. Zhu, L., Shi, G., Schmidt, C. M., Hruban, R. H. & Konieczny, S. F. Acinar cells contribute to the molecular heterogeneity of pancreatic intraepithelial neoplasia. *Am. J. Pathol.* **171**, 263–273 (2007).
10. Guerra, C. *et al.* Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. *Cancer Cell* **11**, 291–302 (2007).
11. Means, A. L. *et al.* Pancreatic epithelial plasticity mediated by acinar cell transdifferentiation and generation of nestin-positive intermediates. *Development* **132**, 3767–3776 (2005).
12. Morris, J. P. t., Cano, D. A., Sekine, S., Wang, S. C. & Hebrok, M.  $\beta$ -catenin blocks Kras-dependent reprogramming of acini into pancreatic cancer precursor lesions in mice. *J. Clin. Invest.* **120**, 508–520 (2010).
13. Bardeesy, N. *et al.* Both p16<sup>ink4a</sup> and the p19<sup>Arf</sup>-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. *Proc. Natl Acad. Sci. USA* **103**, 5947–5952 (2006).
14. Feldser, D. M. *et al.* Stage-specific sensitivity to p53 restoration during lung cancer progression. *Nature* **468**, 572–575 (2010).
15. Junttila, M. R. *et al.* Selective activation of p53-mediated tumour suppression in high-grade tumours. *Nature* **468**, 567–571 (2010).
16. Johnson, L. *et al.* Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116 (2001).
17. Levy, D. *et al.* A proteomic approach for the identification of novel lysine methyltransferase substrates. *Epigenetics Chromatin* **4**, 19 (2011).
18. Maruyama, T. *et al.* CHIP-dependent termination of MEKK2 regulates temporal ERK activation required for proper hyperosmotic response. *EMBO J.* **29**, 2501–2514 (2010).
19. Sun, W. *et al.* MEK kinase 2 and the adaptor protein Lad regulate extracellular signal-regulated kinase 5 activation by epidermal growth factor via Src. *Mol. Cell Biol.* **23**, 2298–2308 (2003).
20. Fanger, G. R., Johnson, N. L. & Johnson, G. L. MEK kinases are regulated by EGF and selectively interact with Rac/Cdc42. *EMBO J.* **16**, 4961–4972 (1997).
21. Choi, M. C. *et al.* A direct HDAC4-MAP kinase crosstalk activates muscle atrophy program. *Mol. Cell* **47**, 122–132 (2012).
22. Matitau, A. E., Gabor, T. V., Gill, R. M. & Scheid, M. P. MEKK2 kinase association with 14-3-3 protein regulates activation of c-Jun N-terminal kinase. *J. Biol. Chem.* **288**, 28293–28302 (2013).
23. Enomoto, A. *et al.* Negative regulation of MEKK1/2 signaling by serine-threonine kinase 38 (STK38). *Oncogene* **27**, 1930–1938 (2008).
24. Brunet, A., Pages, G. & Pouyssegur, J. Constitutively active mutants of MAP kinase kinase (MEK1) induce growth factor-relaxation and oncogenicity when expressed in fibroblasts. *Oncogene* **9**, 3379–3387 (1994).
25. Raman, M., Chen, W. & Cobb, M. H. Differential regulation and properties of MAPKs. *Oncogene* **26**, 3100–3112 (2007).
26. Eichhorn, P. J., Creighton, M. P. & Bernards, R. Protein phosphatase 2A regulatory subunits and cancer. *Biochim. Biophys. Acta* **1795**, 1–15 (2009).
27. Li, Y. M. & Casida, J. E. Cantharidin-binding protein: identification as protein phosphatase 2A. *Proc. Natl Acad. Sci. USA* **89**, 11867–11870 (1992).

**Acknowledgements** We thank members of the Gozani and Sage laboratories for critical reading of the manuscript, A. Smits for help with mass spectrometry data visualization and P.J. Utz and the Floren Family Trust for providing ProtoArrays. This work was supported in part by grants from the NIH to O.G. and J.S. (R01 CA172560) and an NIH Innovator grant (DP2 OD007447) from the Office of the Director for B.A.G.; M.V. was supported by a grant from NWO-VIDI, P.K.M. was supported by the Tobacco-Related Disease Research Program, a Dean's Fellowship from Stanford University, and the Child Health Research Institute and Lucile Packard Foundation for Children's Health at Stanford. N.R. was supported by a grant from the Fondation pour la Recherche Médicale. J.S. is the Harriet and Mary Zelencik Scientist in Children's Cancer and Blood Diseases.

**Author Contributions** N.R. and P.K.M. contributed equally to this work and are listed alphabetically. They were responsible for the experimental design, execution, data analysis and manuscript preparation. P.K. and A.J.B. performed the bioinformatics meta-analysis. P.W.T.C.J. and M.V. performed the SILAC experiments. S.L. and B.A.G. performed the methylated peptide mass spectrometry experiments. A.W.W. generated recombinant H3 and H3K4R protein. O.B., G.S.V., M.H., D.D., P.J.T. and R.G.K. generated SMYD3 and MAP3K2me antibodies, the MAP3K2 peptides, and determined the catalytic efficiency of SMYD3. O.G. and J.S. were equally responsible for supervision of research, data interpretation and manuscript preparation.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.S. ([julsage@stanford.edu](mailto:julsage@stanford.edu)) and O.G. ([ogozani@stanford.edu](mailto:ogozani@stanford.edu)).



## METHODS

**Ethics statement.** Mice were maintained according to practices prescribed by the NIH at Stanford's Research Animal Facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care.

**Mouse strains.** *Kras*<sup>+/LSL-G12D</sup>, *p53*<sup>lox/lox</sup>, and *p48*<sup>Cre/+</sup> mice have been described before<sup>8,28,29</sup>. *Smyd3*<sup>tm1a(KOMP)Wtsi</sup> mice were obtained from the KOMP Repository. Details on the targeted allele are available on the KOMP website<sup>6</sup>. Briefly, mice were constructed using the 'knockout first' strategy. In this allele, insertion of a *LacZ* cassette with strong splice acceptor in intron 2 of the *Smyd3* gene creates a knockout allele serving additionally as a reporter. Expression of the Cre recombinase in cells removes the *LacZ* cassette and further deletes several *Smyd3* exons, resulting in a null allele (Extended data Fig. 2d). Mice were of mixed C57BL/6;129SV background and we systematically used randomly picked littermates as controls in all the experiments (sex ratio per cohort balanced).

**Pancreatic cancer mouse models.** *Pancreatitis-induced tumorigenesis.* Acute pancreatitis was induced at 6 to 8 weeks of age in *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup> (*Kras*) and *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup>; *Smyd3*<sup>lox/lox</sup> (*Kras*; *Smyd3*) mice by administration of 8 hourly intraperitoneal injections of caerulein (125 µg per kg body weight), (Sigma-Aldrich) over 2 consecutive days as described previously<sup>12</sup>. Mice were treated as indicated with the PP2A inhibitor cantharidin (0.15 mg kg<sup>-1</sup> BID, IP), (Sigma-Aldrich) or the MEK inhibitor trametinib (Selleckchem) (1 mg kg<sup>-1</sup> or 0.1 mg kg<sup>-1</sup> intraperitoneally daily) or vehicle 10% cyclodextran. Pancreatic lesions were analysed 7 days after the last injection.

**Spontaneous model of pancreatic intraepithelial neoplasia (PanIN) development.** PanIN progression was analysed in *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup> (*Kras*) and *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup>; *Smyd3*<sup>lox/lox</sup> (*Kras*; *Smyd3*) mice aged for 6 months. Quantification of low (PanIN1a and 1b) and high-grade (PanIN2 and 3) lesions was performed. Histopathological analysis was conducted on de-identified slides based on the classification consensus<sup>30</sup>. Five images (×100) were taken in standardized positions (as to cover the whole section) for each slide. PanINs were counted from 8 independent animals for each group. Error bars represent s.e.m.

**Model of aggressive PDAC.** To study aggressive PDAC expansion, we generated *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup>; *p53*<sup>lox/lox</sup> (*Kras*; *p53*) and *p48*<sup>Cre/+</sup>; *Kras*<sup>LSL-G12D</sup>; *p53*<sup>lox/lox</sup>; *Smyd3*<sup>lox/lox</sup> (*Kras*; *p53*; *Smyd3*) mutant mice. Mice were followed for signs of disease progression. At endpoint, tumours were processed for histological and immunohistochemical evaluation. To calculate relative normal acini area *Kras*; *p53* and *Kras*; *p53*; *Smyd3* tumour sections were stained for amylase. Positive regions on six random, non-overlapping, ×100 images were collected from 3 mice per genotype. For each image positive amylase area was normalized to total pancreas tissue area using ImageJ software. Error bars represent means ± standard error of the mean (s.e.m.).

### Lung cancer mouse models

**Adenovirus-induced lung adenocarcinoma (LAC).** *Kras*<sup>LSL-G12D</sup> (*Kras*) and *Kras*<sup>LSL-G12D</sup>; *Smyd3*<sup>lox/lox</sup> (*Kras*; *Smyd3*) mice were treated with 5 × 10<sup>6</sup> plaque-forming units of adenovirus expressing Cre (University of Iowa adenovirus core) by intratracheal infection as previously described<sup>31</sup>. Tumours were analysed and quantified at 12, 16 and 20 weeks post-infection, *n* = 6 for each group.

**Lentivirus-induced lung adenocarcinoma.** Generation of a dual promoter lentiviral vector for Cre and cDNA expression was described before<sup>32</sup>. To determine the effect of reconstitution of exogenous *Smyd3*<sup>WT</sup> and *Smyd3*<sup>F183A</sup> expression on lung tumour progression a lentiviral vector was developed that expressed both the *Smyd3* complementary DNA and Cre. A lentivirus expressing Cre alone was used as a control. Virus was produced and titred as described previously<sup>33</sup>. Briefly, the lenti-Cre vector was co-transfected with packaging vectors into 293T cells using calcium-phosphate. The supernatant was collected at 48 and 72 h. Concentrated virus was recovered by ultracentrifugation at 25,000 r.p.m. for 90 min and resuspended in PBS. Cohorts of *Kras*<sup>LSL-G12D</sup>; *Smyd3*<sup>lox/lox</sup> (*Kras*; *Smyd3*) mice were infected with each lentiviral vector. Tumour burden was analysed 24 weeks after lentiviral infection, *n* = 4 for each treatment.

**Preparation of pancreatic epithelial explants culture.** Pancreatic epithelial explants from 4- to 6-week-old *p48*<sup>Cre/+</sup> (WT) and *p48*<sup>Cre/+</sup>; *Smyd3*<sup>lox/lox</sup> (*Smyd3*) were established by modification of previously published protocols<sup>34</sup>. In brief, the whole pancreas was collected and treated twice with 1.2 mg ml<sup>-1</sup> collagenase VIII (Sigma-Aldrich). Following multiple wash steps with McCoy's medium containing soybean trypsin inhibitor (SBTI, 0.2 mg ml<sup>-1</sup>), digested samples were filtered through a 100-µm filter, resuspended in culture medium (Waymouth's MB 752/1 supplemented with 0.1% BSA, 0.2 mg ml<sup>-1</sup> SBTI; 50 µg ml<sup>-1</sup> bovine pituitary extract, 10 µg ml<sup>-1</sup> insulin, 5 µg ml<sup>-1</sup> transferrin, 6.7 ng ml<sup>-1</sup> selenium in 30% FCS) and allowed to recover for 1 h at 37 °C. Thereafter, cells were pelleted and resuspended in culture medium supplemented with penicillin G (1,000 U ml<sup>-1</sup>), streptomycin (100 µg ml<sup>-1</sup>), amphotericin B (0.25 µg ml<sup>-1</sup>), 0.1% FCS, and an equal volume of rat tail collagen type I (BD Bioscience). The cellular/rat tail collagen type I suspension was immediately plated on plates pre-coated with 2.5 mg ml<sup>-1</sup> of rat tail collagen type I. In stimulation experiments recombinant human EGF (rhEGF, Invitrogen) was added

at final concentration of 25 ng ml<sup>-1</sup>. For quantification, acinar explants were seeded in triplicates. Cells clusters were counted from at least 3 optical fields/well and reported as a percentage of acinar clusters and duct-like spheres. The quantification was performed in two independent experiments; the number of mice is reported in the main text.

**Lentiviral transduction of pancreatic epithelial explants.** For reconstitution experiments fresh explants from *Kras*<sup>LSL-G12D</sup>; *Smyd3*<sup>lacZ/lacZ</sup> (*Smyd3* deficient) were transduced using lentiviral vectors with or without Cre and expressing wild-type SMYD3 or SMYD3(F183A). Cre-mediated recombination results in oncogenic K-Ras activation, causing spontaneous ADM (acinar to ductal metaplasia), which was quantified as described above.

**Immunofluorescence of pancreatic epithelial explants.** For immunofluorescent labelling of explanted pancreatic acini/duct clusters, collagen gels containing explanted pancreas were fixed in the chamber slides in 4:1 methanol:DMSO overnight at 4 °C, washed and stored at -20 °C in 100% methanol. Collagen gels were permeabilized with TritonX-100 0.1% (Sigma-Aldrich) for 5 min at room temperature and washed in PBS+TritonX-100 0.025% (PBST). Gels were blocked with 5% normal goat serum in PBST+BSA 2% for 2 h at room temperature, then incubated sequentially with SMYD3 primary antibody (Abcam) and Alexa Fluor 488-conjugated secondary antibody diluted in blocking buffer overnight at 4 °C. Following each antibody, gels were washed in PBST. Cells were counterstained with 50 µg ml<sup>-1</sup> DAPI (Invitrogen) and washed in PBS. Images were captured on a Zeiss inverse fluorescent microscope. Identical acquisition methods were used for all samples to allow direct comparison of the resulting images.

**Histology, immunohistochemistry and X-gal staining.** Tissue specimens were fixed in 4% buffered formalin for 24 h and stored in 70% ethanol until paraffin embedding. 3-µm sections were stained with haematoxylin and eosin (HE) or used for immunohistochemical studies.

**Immunohistochemistry.** Immunohistochemistry was performed on formalin-fixed, paraffin embedded mouse and human tissue sections using a biotin-avidin method as described before<sup>35</sup>. The following antibodies were used: rabbit anti-Amylase (Sigma-Aldrich), pERK1/2 (Cell Signaling), MUC5 (NeoMarkers) and SMYD3 (Abcam). Sections were developed with DAB and counterstained with haematoxylin. Pictures were taken using a Zeiss microscope equipped with the AxioVision software. Analysis of the tumour area and IHC analysis was done using ImageJ software by measuring pixel units.

**X-gal staining.** Staining of cryosections (8 µm) was carried out as described previously<sup>35</sup>; slides were counterstained with nuclear fast red.

**qRT-PCR.** RNA was isolated using the Qiagen RNeasy Isolation Kit followed by cDNA synthesis (SuperScript II, Invitrogen). Real-time PCR was performed with 800 nM primers diluted in a final volume of 20 µl in SYBR Green Reaction Mix (Applied Biosystems). RT-PCRs were performed as follows: 95 °C for 10 min, 35 cycles of 95 °C for 15 s and 60 °C for 1 min. qRT-PCR data are representative of 4 independent mouse pancreatic epithelial explants isolations per treatment. All samples were analysed in triplicate. *Gapdh* expression was used for normalization. The following primers were used:

*Smyd3*-For 5'-TGCGCACCATTGGAGCCGTAC  
*Smyd3*-Rev 5'-GTCAAAGGCCAGCCTCAGTTCT  
*Gapdh*-For 5'-CCCACTAACATCAATGGGG  
*Gapdh*-Rev 5'-CCTTCCACAATGCCAAAGTT

**Meta-analysis of public PDAC and NSCLC data sets.** We downloaded raw data for gene expression studies (7 pancreatic cancer, 6 NSCLC) from the NCBI GEO and EBI ArrayExpress. After re-annotating the probes, each data set was normalized separately using gcRMA. We applied two meta-analyses approaches to the normalized data. The meta-analysis approach has been recently described<sup>36</sup>. Briefly, the first approach combines effect sizes from each data set into a meta-effect size to estimate the amount of change in expression across all data sets. For each gene in each data set, an effect size was computed using Hedges' adjusted *g*. If multiple probes mapped to a gene, the effect size for each gene was summarized using the fixed effect inverse-variance model. We combined study-specific effect sizes to obtain the pooled effect size and its standard error using the random effects inverse-variance technique. We computed *z*-statistics as a ratio of the pooled effect size to its standard error for each gene, and compared the result to a standard normal distribution to obtain a nominal *P* value. *P* values were corrected for multiple hypotheses testing using false discovery rate (FDR)<sup>37</sup>. We used a second non-parametric meta-analysis that combines *P* values from individual experiments to identify genes with a large effect size in all data sets. Briefly, we calculated a *t*-statistic for each gene in each study. After computing one-tail *P* values for each gene, they were corrected for multiple hypotheses using FDR. Next, we used Fisher's sum of logs method, which sums the logarithm of corrected *P* values across all data sets for each gene, and compares the sum against a chi-square distribution with 2*k* degrees of freedom, where *k* is the number of data sets used in the analysis.



**Plasmids.** Bacterial expression plasmids were created using pGEX-6P1 vector. Transient mammalian expression plasmids were created using pCDNA3.1 HA, pCDNA3.1 Myc and pCagFlag vectors. The different inserts were amplified by PCR using either cDNA or specific clones from the human ORFeome library as template. Single-point mutations of SMYD3 and MAP3K2 were generated using the QuikChange site-directed mutagenesis protocol (Stratagene), and clones were confirmed by DNA sequencing. SMYD3 and MAP3K2 shRNA targeting UTR regions were cloned in a pSICOR vector carrying a puromycin resistance gene.

Human SMYD3 shRNA sequence directed against the 3' UTR TGGTGTG TCTTTGTTGAATTTCAAGAGAATTCACAAAGACACGCTTTTTTC.

Human MAP3K2 shRNA sequence directed against the 3' UTR TGGATGA TTTCAGTATGGCATTTCAGAGAATGCCTAGTGAAATCATCCTTTTTTC.

Mouse *Smad3* shRNA sequence directed against the 3' UTR TGAGCAGAACC ATTACAATTTCAAGAGATATTGTAATGGTTCTGCTCTTTTTTC.

Mouse *Map3k2* shRNA sequence directed against the 3' UTR TAGTCATAGC TATAGTGAAATTCAGAGATTTCACTATAGCTATGACTTTTTTTC. SMYD3 and MAP3K2 stable reconstitution plasmids have been created using the Gateway cloning system according to the manufacturer's instructions (Invitrogen) with either the WT or point mutant constructs, into pMSCV-Flag and pBabe-Flag vectors (hygromycin resistance).

**Cell culture, reagents and transfections.** K-Ras mutant lung and pancreatic carcinoma lines LKR10 (mouse lung), A549 (human lung) and CFPac1 (human pancreas) were used (all these cell lines are wild-type for EGFR, see below and the COSMIC database for the human cell lines). 293T, LKR10, A549 and CFPac1 cells were grown in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 10% fetal calf serum (FCS, GIBCO), 100 units ml<sup>-1</sup> penicillin/streptomycin and glutamine. All cells were cultured at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. For transient expression, cells were transfected using Mirus 293T transfection reagent and collected 24 to 36 h later. For stable knockdown, cells were transduced with lentiviral shRNA constructs using the packaging vectors pGagpol and pΔ8.2, followed by 2 μg ml<sup>-1</sup> puromycin selection for one week. For rescue experiments, cells were transduced with retroviral pBabe and pMSCV constructs using packaging the pGag and pVSVg vectors, followed by 100 μg ml<sup>-1</sup> hygromycin selection for one week.

Serum and EGF stimulation were performed after 48 h of serum-starvation using either regular 10% FBS cell media or 25 ng μl<sup>-1</sup> of rhEGF (Promega) for 15 min. **EGFR sequencing.** LKR10 DNA was isolated. *Egfr* exons 18 to 21 were amplified using a proofreading polymerase (PfuUltra, Invitrogen) and sequenced bi-directionally. Sequence was verified based on transcript ID: ENSMUST00000020329. The following primers were used:

*mEGFR Exon 18:* For-5' CTCTGGCTCAGAATGAATCTAC, Rev-5' GAAGCCT AGTGCGGACCTGTC, product: 268 bp.

*mEGFR Exon 19:* For-5' CCAGCTCACAAGGCAACATG, Rev-5' CTAAGGAAG CAAGATTGACC, product: 229 bp.

*mEGFR Exon 20:* For-5' GATTCATCTATTGTCCTTACC, Rev-5' TGGGTACT TCAGTGGACAGAC, product: 234 bp.

*mEGFR Exon 21:* For-5' CATGACACTGAGGATGCCCAGA, Rev-5' CAAATGC TGCCACAGCTGAC, product: 298 bp.

**Cell extracts, immunoprecipitation and immunoblot analysis.** For total cell extracts, cells were lysed in RIPA buffer (10 mM Tris-HCl pH 8, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton, 0.1% SDS, 1 mM PMSF, protease inhibitors (Roche) and a phosphatase inhibitor cocktail (Sigma-Aldrich)) for 15 min. Cell fractionation was performed by collecting supernatant (cytoplasmic fraction) after 5 min of 1,300g centrifugation following a 10 min incubation in hypotonic buffer (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.34 M sucrose, 10% glycerol, 1 mM DTT, 0.05% Triton, and protease inhibitors). The pellet was then incubated 15 min in LSDB250 buffer (glycerol 20%, MgCl<sub>2</sub> 3 mM, HEPES pH 7.9 50 mM, KCl 250 mM, DTT 0.5 mM, PMSF 0.5 mM, NP40 0.1%, protease inhibitors), and centrifuge at 15,000g for 10 min. The supernatant was collected as soluble nuclear extract, and the pellet was further extracted in LSDB250 buffer with sonication (chromatin fraction). Protein concentration was determined by the BCA assay (Pierce).

For immunoprecipitation, cells were lysed in either LSDB250 buffer for smooth total cell extract, or hypotonic buffer for cytoplasmic extract, and same amount of protein extracts were incubated with specific antibody overnight at 4 °C. Extracts were then incubated with protein A Sepharose beads (GE Healthcare) for 2 h at 4 °C.

Proteins were resolved by SDS-PAGE, transferred to nitrocellulose membrane and analysed by immunoblot. Antibodies used were as follows: SMYD3 and MAP3K2me2/3 (generated by Yenzym); GST (generated by Covance); beta-tubulin (05-661 Millipore); MAP3K2 (1662-1 Epitomics); Flag and ERK5 (F1804, E1523, Sigma); pERK1/2 and ERK1/2 and pMEK1/2 and MEK1/2 and pJNK and JNK and pP38 and P38 and pERK5 and pAKT and AKT and pCraf and CRaf, PPP2R2a (4370, 4695, 9121, 9122, 9251, 9252, 9211, 9212, 3371, 2965, 4685, 9427, 9422, 5689, Cell Signaling); HA

and Myc (26183, 21316 Pierce). Immunoblots signal intensity was measured using ImageJ software. Quantification data are expressed as s.e.m.

**Expression and purification of recombinant proteins.** For expression of GST tagged recombinant proteins, transformed BL21 cells were induced with 0.1 mM IPTG overnight at 20 °C, and proteins were purified using glutathione Sepharose beads (GE Healthcare) and eluted in 10 mM reduced glutathione (Sigma) or cleaved from the GST tag using purified Precision enzyme. Recombinant human histone H3 and H3K4\* mutant (first thirteen lysines mutated to arginine except lysine K4) were expressed and purified as previously described<sup>38</sup>.

**ProtoArray, methylation and kinase assays.** *In vitro* methylation assays were performed using 1 to 2 μg of recombinant proteins or peptides incubated with 1 μg of recombinant methyltransferases and 0.1 mM of S-adenosyl-methionine (SAM, Sigma) or 2 μCi <sup>3</sup>H-AdoMet (American Radiolabelled Chemicals) in buffer containing 50 mM Tris-HCl (pH 8.0), 10% glycerol, 20 mM KCl, 5 mM MgCl<sub>2</sub>, and 1 mM PMSF at 30 °C overnight. The reaction mixture was resolved by SDS-PAGE, followed by autoradiography, Coomassie stain (Pierce) or mass spectrometry analysis.

Kinetic constants *k*<sub>cat</sub> and *K*<sub>M</sub> were determined using Grafit (Erithacus software) from SMYD3 methylation activity on histone H4 (EMD Millipore, USA) and MAP3K2 (Origene, USA) was assessed by radiometric assays using [<sup>3</sup>H]SAM with specific product capture on arginine binding SPA beads for assays using histone substrates or RNA binding SPA beads when MAP3K2 was the substrate. (PerkinElmer, USA). Assay conditions were 20 mM Tris pH 8, 3 mM DTT, 50 μM ZnCl<sub>2</sub>, 0.005% Tween-20, 1.2 μM unlabelled SAM, 0.2 μM [<sup>3</sup>H]SAM and 25 nM SMYD3 final. Reactions were quenched using 2 mM unlabelled SAM. SPA signal was quantified in a Microbeta scintillation counter (PerkinElmer, USA). Data were fit to the Michaelis-Menten equation<sup>6</sup>, where the rate was plotted as a function of the concentration of substrate.

$$v = \frac{V_{\max}[S]}{K_M + [S]}$$

To verify the activity of the other lysine methyltransferases tested, *in vitro* methylation assays were performed using known substrates, as previously reported<sup>39-48</sup>.

*In vitro* kinase assays were performed by incubating 1 μg of recombinant MAP3K2 WT or mutants with 1 μg of recombinant MEK1 in kinase buffer containing 25 mM Tris-HCl (pH 7.5), 5 mM β-glycerophosphate, 2 mM dithiothreitol (DTT), 0.1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 mM MgCl<sub>2</sub> and 200 μM ATP (Cell Signaling) at 37 °C during 30 min. **Peptide pull-down and SILAC.** MAP3K2 peptides were generated by 21st Century Biomaterials and are based on the following sequence: DYDNPIFEKFGKGGTYP RRYHVSYHG[K-Biot-Ahx]-amide.

For peptide pull-downs, 5 to 10 μl of streptavidin Sepharose beads (GE Healthcare) were saturated with 10 μg of specific biotinylated peptides for 2 h at 4 °C under rotation in peptide buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% NP-40), then washed 3 times in the same buffer. Beads were then incubated with either 1 μg of recombinant proteins or 1 mg of HeLa cytoplasmic extract for 4 h at 4 °C under rotation in peptide buffer. Beads were then washed 3 times in the same buffer and resuspended in Laemmli buffer for immunoblot analysis.

For SILAC peptide pull-down, HeLa cytoplasmic extracts were prepared from cells cultivated in either normal amino acids culture condition ('light') or using modified amino acids culture condition ('heavy'). A 2-way experiments was performed, the 'forward' condition combining MAP3K2-K260me0 peptide with light extract and MAP3K2-K260me3 peptide with heavy extract, the 'reverse' condition combining MAP3K2-K260me3 peptide with light extract and MAP3K2-K260me0 peptide with heavy extract. Beads of each pair of peptide pull-down were then pooled together, washed and extracts were resuspended in Laemmli buffer and resolved by SDS-PAGE. In-gel trypsin digestion was performed and peptides purified using C18 stage tips (Fisher) before mass spectrometry analysis to quantify the ratio of each potential binders for K260me0 and K260me3 peptides, in forward and reverse condition. To identify outliers in both the forward and the reverse experiment, boxplot statistics was applied (cut off = 1.5 × interquartile range). Proteins identified as outlier in both experiments are assigned as significant interactors. Amino acid complements used for SILAC are L-lysine-2HCl (Thermo Scientific 88429), L-arginine-HCl (Thermo Scientific 88427), <sup>2</sup>H<sub>4</sub>-L-lysine-2HCl (Thermo Scientific 88438), <sup>13</sup>C<sub>6</sub>-L-arginine-HCl (Thermo Scientific 88433), L-proline (Thermo Scientific 88430).

**Active Ras pull-down and detection.** Ras activity refers to the level of guanosine triphosphate-bound Ras, which is able to bind Ras binding domain (RBD) of RAF-1 as measured using a RBD-domain pull-down assay kit as recommended by the manufacturer (The Active Ras Pull-Down and Detection Kit, Thermo Scientific). Briefly, tumour biopsies were homogenized on ice in lysis buffer containing 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (pH 7.5), 1% Igepal CA-630, 150 mM NaCl, 0.25% sodium deoxycholate, 10% glycerol, 25 mM NaF, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, 10 μg ml<sup>-1</sup> aprotinin, 10 μg ml<sup>-1</sup> leupeptin and 1 mM sodium orthovanadate. These samples were sonicated and centrifuged at 15,000g for 10 min at 4 °C to remove cellular debris. Protein concentration was measured. Equal amounts

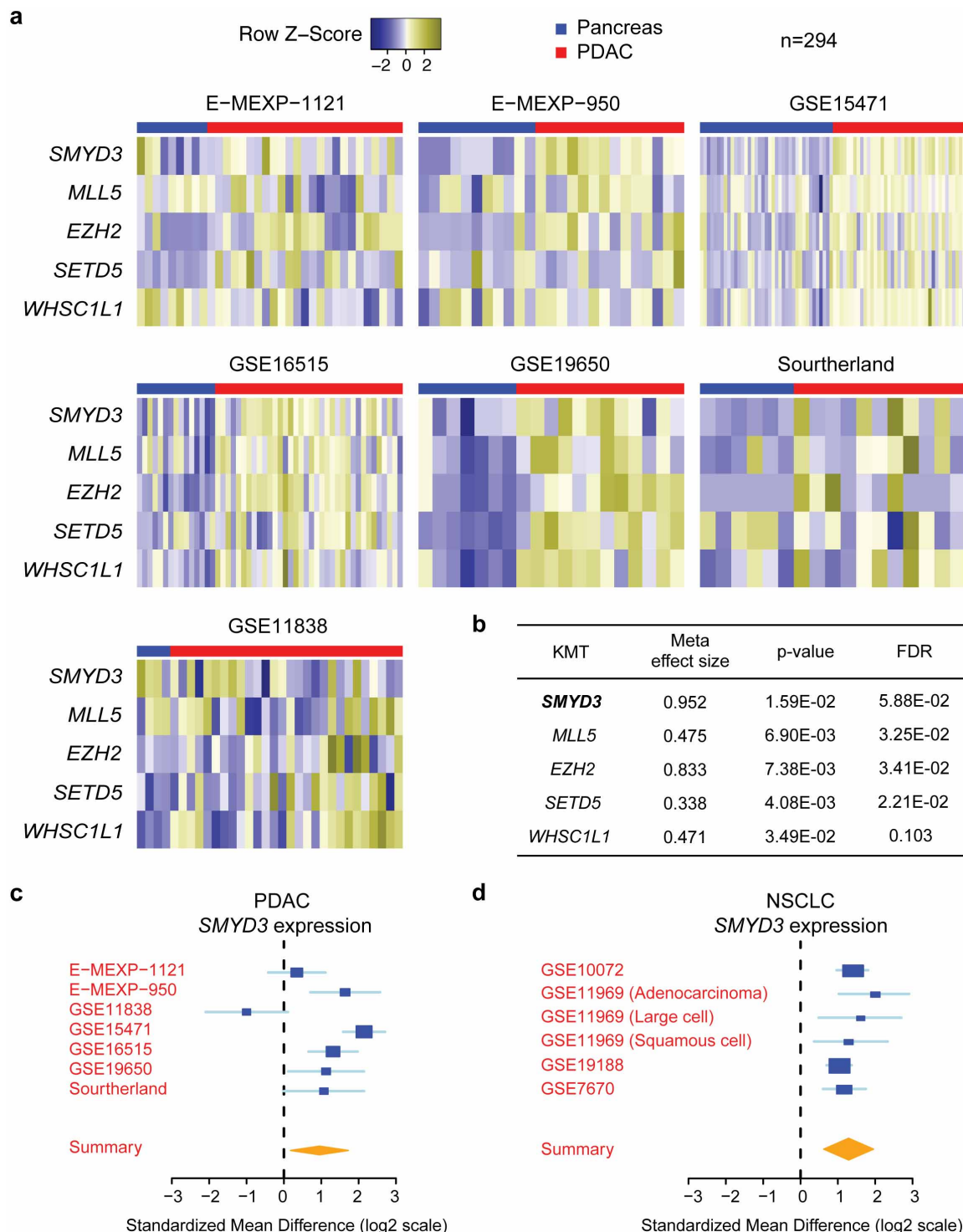
of lysate were incubated for 30 min at 4 °C with agarose beads coated with RBD. The beads were then washed three times with ice-cold lysis buffer, boiled for 5 min at 95 °C, and active Ras was analysed by immunoblotting following standard protocol using Ras-specific antibodies (Thermo Scientific). For comparison to total Ras protein, 2% of total lysates used for pull-down was analysed by immunoblot.

**Cancer xenografts.** For xenograft analysis, 500,000 CFPac1 cells were injected into the flank of NSG mice with Matrigel (BD Bioscience). Tumour volume was measured at the times indicated and calculated using the ellipsoid formula (length × width<sup>2</sup>).

**Cell assays.** Anchorage-independent growth was assessed in soft agar assays. Cell proliferation was assessed by counting cell number at indicated days and expressed relative to the control as previously described<sup>6</sup>. Cell viability in response to treatment with the MEK inhibitor trametinib (Selleckchem) was measured by an MTT assay (Roche) according to the manufacturer's instructions.

**Statistics.** Kaplan–Meier survival curves were calculated using the survival time for each mouse from all littermate groups. The log-rank test was used to test for significant differences between the groups. For image quantification and gene expression analysis statistical significance was assayed by Student's *t*-test with the Prism GraphPad software (two-tailed unpaired and paired *t*-test depending on the experiment; variance was first systematically examined using an F-test). \**P* value < 0.05; \*\**P* value < 0.01; \*\*\**P* value < 0.001; ns, not significant. Data are represented as mean ± standard error of the mean (s.e.m.).

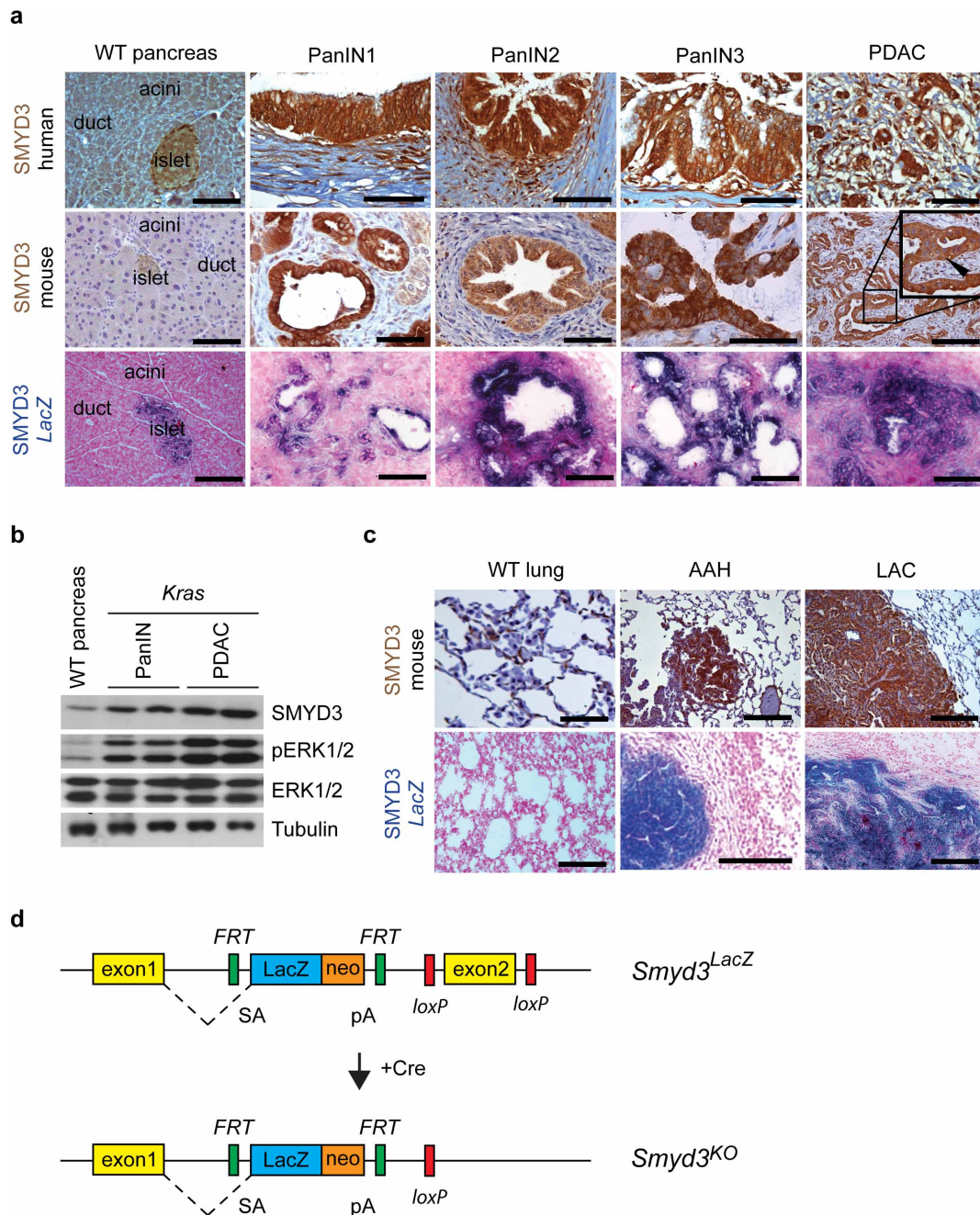
28. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nature Genet.* **29**, 418–425 (2001).
29. Kawaguchi, Y. *et al.* The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nature Genet.* **32**, 128–134 (2002).
30. Hruban, R. H. *et al.* Pathology of genetically engineered mouse models of pancreatic exocrine cancer: consensus report and recommendations. *Cancer Res.* **66**, 95–106 (2006).
31. DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols* **4**, 1064–1072 (2009).
32. Winslow, M. M. *et al.* Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* **473**, 101–104 (2011).
33. Tiscornia, G., Singer, O. & Verma, I. M. Production and purification of lentiviral vectors. *Nature Protocols* **1**, 241–245 (2006).
34. Heid, I. *et al.* Early requirement of Rac1 in a mouse model of pancreatic cancer. *Gastroenterology* **141**, 719–730 (2011).
35. Mazur, P. K. *et al.* Notch2 is required for progression of pancreatic intraepithelial neoplasia and development of pancreatic ductal adenocarcinoma. *Proc. Natl Acad. Sci. USA* **107**, 13438–13443 (2010).
36. Khatri, P. *et al.* A common rejection module (CRM) for acute rejection across multiple organs identifies novel therapeutics for organ transplantation. *J. Exp. Med.* **210**, 2205–2221 (2013).
37. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc. Series B* **64**, 479–498 (2002).
38. Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375**, 23–44 (2003).
39. Brown, M. A., Sims, R. J. III, Gottlieb, P. D. & Tucker, P. W. Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol. Cancer* **5**, 26 (2006).
40. Green, E. M., Mas, G., Young, N. L., Garcia, B. A. & Gozani, O. Methylation of H4 lysines 5, 8 and 12 by yeast Set5 calibrates chromatin stress responses. *Nature Struct. Mol. Biol.* **19**, 361–363 (2012).
41. Grey, C. *et al.* Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. *PLoS Biol.* **9**, e1001176 (2011).
42. Kuo, A. J. *et al.* NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* **44**, 609–620 (2011).
43. Levy, D. *et al.* Lysine methylation of the NF-κB subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF-κB signaling. *Nature Immunol.* **12**, 29–36 (2011).
44. Moore, K. E. *et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol. Cell* **50**, 444–456 (2013).
45. Saddic, L. A. *et al.* Methylation of the retinoblastoma tumor suppressor by SMYD2. *J. Biol. Chem.* **285**, 37733–37740 (2010).
46. Shi, X. *et al.* Modulation of p53 function by SET8-mediated methylation at lysine 382. *Mol. Cell* **27**, 636–646 (2007).
47. Strahl, B. D. *et al.* Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol. Cell Biol.* **22**, 1298–1306 (2002).
48. Tachibana, M., Sugimoto, K., Fukushima, T. & Shinkai, Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* **276**, 25309–25317 (2001).



**Extended Data Figure 1 | SMYD3 is a highly overexpressed KMT in Ras-associated cancers.** **a**, Analysis of seven publicly available human pancreatic ductal adenocarcinoma (PDAC) gene expression studies from the NCBI GEO and EBI ArrayExpress for *SMYD3* levels. The red line indicates expression of *SMYD3* in pancreatic cancer biopsies ( $n = 203$ ); the blue line marks normal pancreas samples ( $n = 91$ ). The scale shows relative expression levels (log<sub>2</sub>). **b**, A bioinformatics meta-analysis identified 5 lysine

methyltransferase overexpressed in human pancreatic ductal adenocarcinoma (PDAC). Meta-effect size and statistical tools are described in the Methods. FDR, false discovery rate. **c**, **d**, Summary of *SMYD3* expression levels in seven ( $n = 294$  independent samples) publicly-available expression data sets of PDAC and six data sets ( $n = 319$  tumours and  $n = 147$  normal independent samples) of non-small cell lung cancer (NSCLC), respectively. Detailed statistical description in the Methods section.

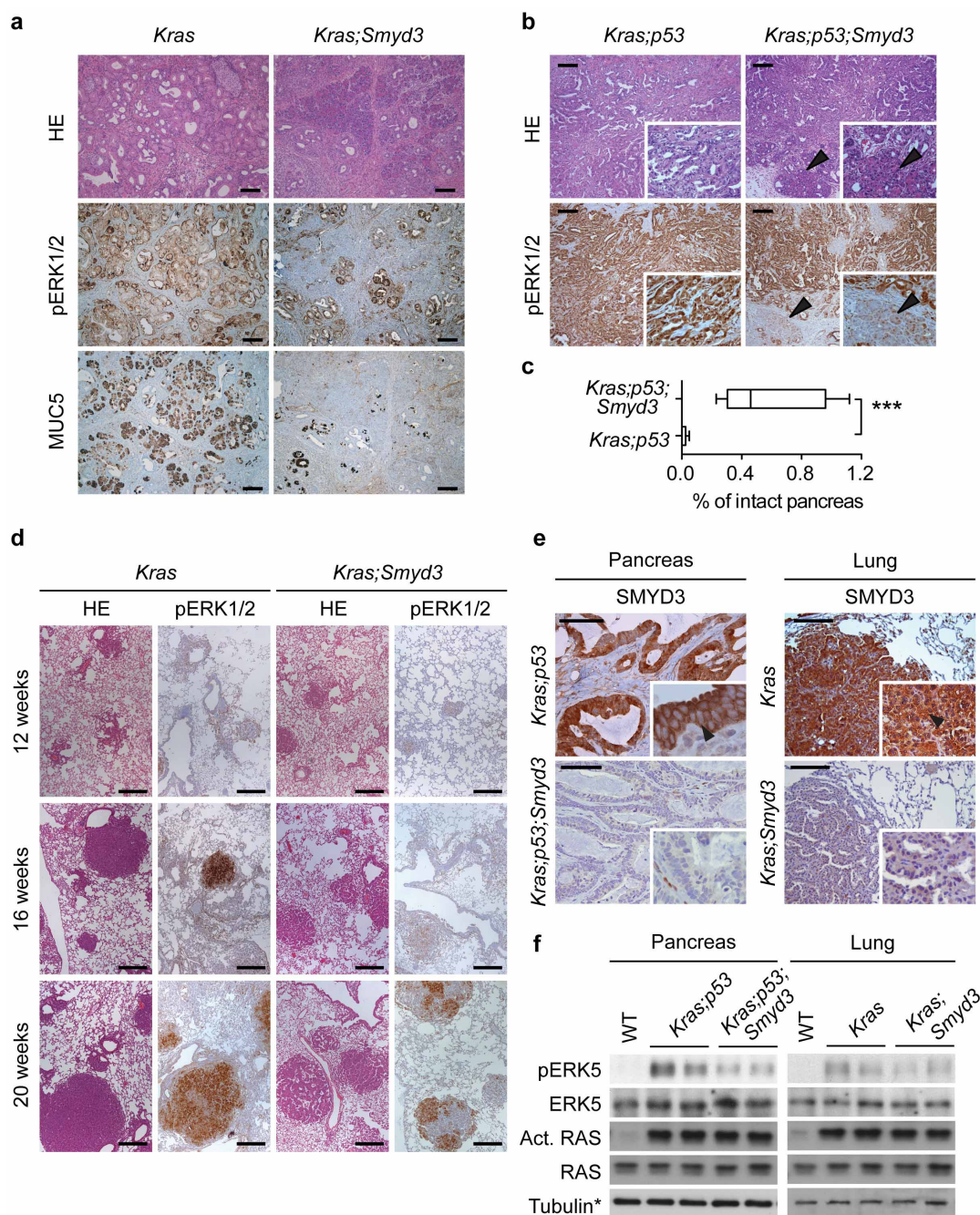




**Extended Data Figure 2 | Analysis of SMYD3 expression in human and mouse PDAC and lung adenocarcinoma (LAC).** **a**, Immunohistochemical analysis of SMYD3 expression in mouse and human WT pancreas, PanIN lesions, and PDAC. The expression pattern was further analysed using a *Smyd3<sup>LacZ</sup>* reporter knock-in strain. *Smyd3<sup>LacZ</sup>* mice were crossed to *p48;Kras<sup>G12D</sup>* (*Kras*) mice and studied at progressing stages of disease. Analysis of *LacZ* activity by X-gal staining as a surrogate for *Smyd3* expression is shown (lower panel) (see Extended Data Fig. 3 for a cartoon of the knock-in allele). **b**, Immunoblot analysis with the indicated antibodies on tumour biopsy lysates from wild-type pancreas and from the pancreas of *Kras* mutant mice at 4.5 and 9 months of age when mice develop PanIN and PDAC, respectively (each

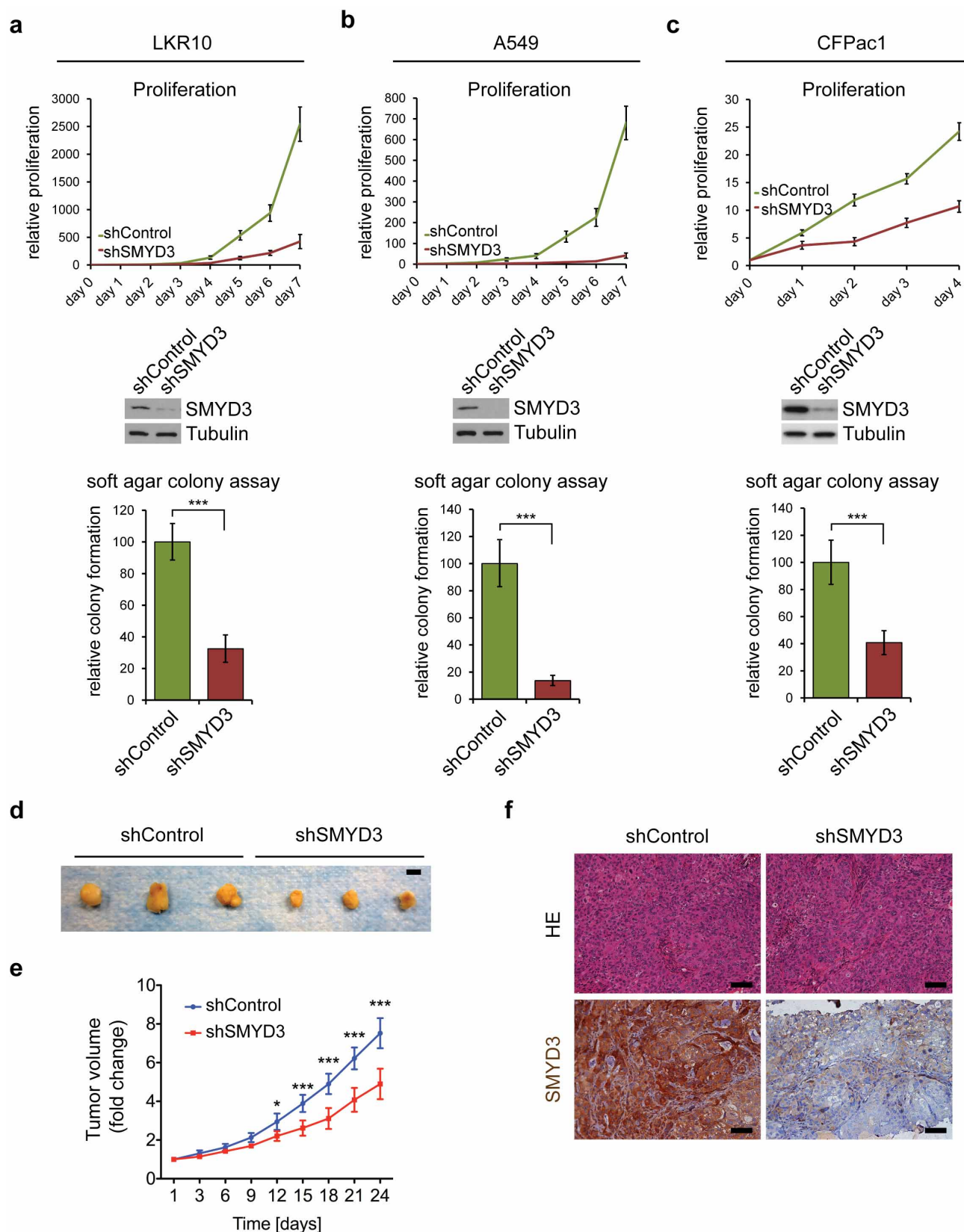
time point represents two biological replicates). **c**, IHC analysis of SMYD3 expression in normal lung, atypical adenomatous hyperplasia (AAH), and lung adenocarcinoma (LAC). X-gal analysis of *LacZ* activity in *Kras*-driven tumours with the *Smyd3<sup>LacZ</sup>* reporter strain (lower panel). All images shown are representative. Arrowheads indicate nuclear localization of SMYD3. Scale bars, 50  $\mu$ m. **d**, *Smyd3* knockout allele diagram. In this allele, insertion of a *LacZ* cassette with a strong splice acceptor in intron 2 of the *Smyd3* gene creates a mutant allele serving additionally as a reporter (*Smyd3<sup>LacZ</sup>*). Expression of the Cre recombinase in cells removes the *LacZ* cassette and further deletes *Smyd3* exon 2, resulting in a null allele *Smyd3<sup>KO</sup>*. SA, splice acceptor; pA, polyadenylation signal.





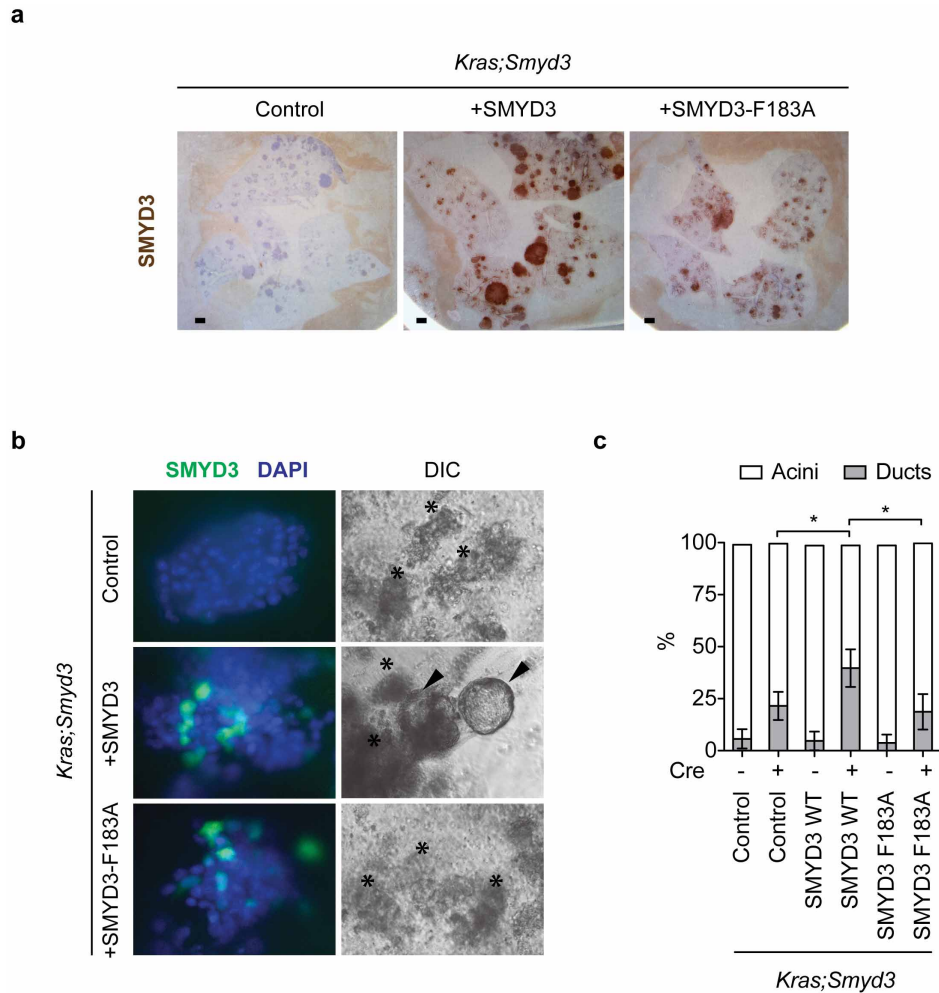
**Extended Data Figure 3 | *Smyd3* deletion inhibits pancreatic tumorigenesis.** **a**, Analysis of pancreatic tumorigenesis at 6 months in *Kras* and *Kras;Smyd3* mutant mice. Representative serial histology section (HE), IHC for pERK1/2 and the PanIN marker MUC5. **b**, Pancreatic cancer phenotypes in *Kras;p53* and *Kras;p53;Smyd3* mutant mice. Representative IHC for pERK1/2. Arrowheads indicate areas with intact acinar cells. **c**, Quantification of intact normal acinar area (amylase-positive area) in *Kras;p53* and *Kras;p53;Smyd3* mutant mice. Data are represented as mean  $\pm$  s.e.m. \*\*\**P* value < 0.001 (two-tailed unpaired Student's *t*-test). **d**, Representative HE and pERK1/2 IHC images of lung sections from *Kras*

and *Kras;Smyd3* mutant mice 12, 16 and 20 weeks after Ad-Cre infection. pERK1/2 is a marker of Ras activity and advanced tumours. **e**, **f**, IHC analysis of SMYD3 expression in the PDAC (**e**) and LAC (**f**) mouse models. Arrowheads indicate cytoplasmic localization of SMYD3. Scale bars, 50  $\mu$ m. **g**, **h**, Immunoblot analysis with the indicated antibodies probing pancreatic adenocarcinoma (**g**) or lung adenocarcinoma tumour lysates (**h**) dissected from *Kras* and *Kras;Smyd3* mutant mice. Active Ras corresponds to Ras protein in the GTP-bound state pulled down with the RAF Ras-binding domain (RBD) (see Methods). \*Tubulin loading control as in Fig. 1j and 2f, respectively.



**Extended Data Figure 4 | SMYD3 functions to maintain the tumorigenic characteristics of human and murine cancer cells.** **a–c**, Cell proliferation rates (top panels) and colony formation in soft agar assays (bottom panels) of murine LAC cell line LKR10 (**a**), human LAC cell line A549 (**b**), or human PDAC cell line CFPac1 (**c**) with or without SMYD3 depletion by stable shRNA (respective immunoblot in middle panels). **d–f**, SMYD3 depletion in CFPac1 attenuates tumour growth in mouse xenografts. **d**, Macroscopic picture of

xenografts from control and SMYD3 knock-down tumours at the end of the experiment. Scale bar, 1 cm. **e**, Volume analysis shows that shSMYD3 significantly inhibits the expansion of pancreatic tumours ( $n = 6$  for each group). **f**, HE of the tumours and IHC confirmation of SMYD3 expression and knock-down. All scale bars, 50  $\mu$ m. \* $P$  value < 0.05; \*\* $P$  value < 0.01; \*\*\* $P$  value < 0.001 (two-tailed unpaired Student's  $t$ -test). Data are represented as mean  $\pm$  s.e.m.



**Extended Data Figure 5 | Lentiviral reconstitution of SMYD3 in pancreatic acinar-to-ductal-metaplasia (ADM) assays and in lung cancer cells *in vivo*.**

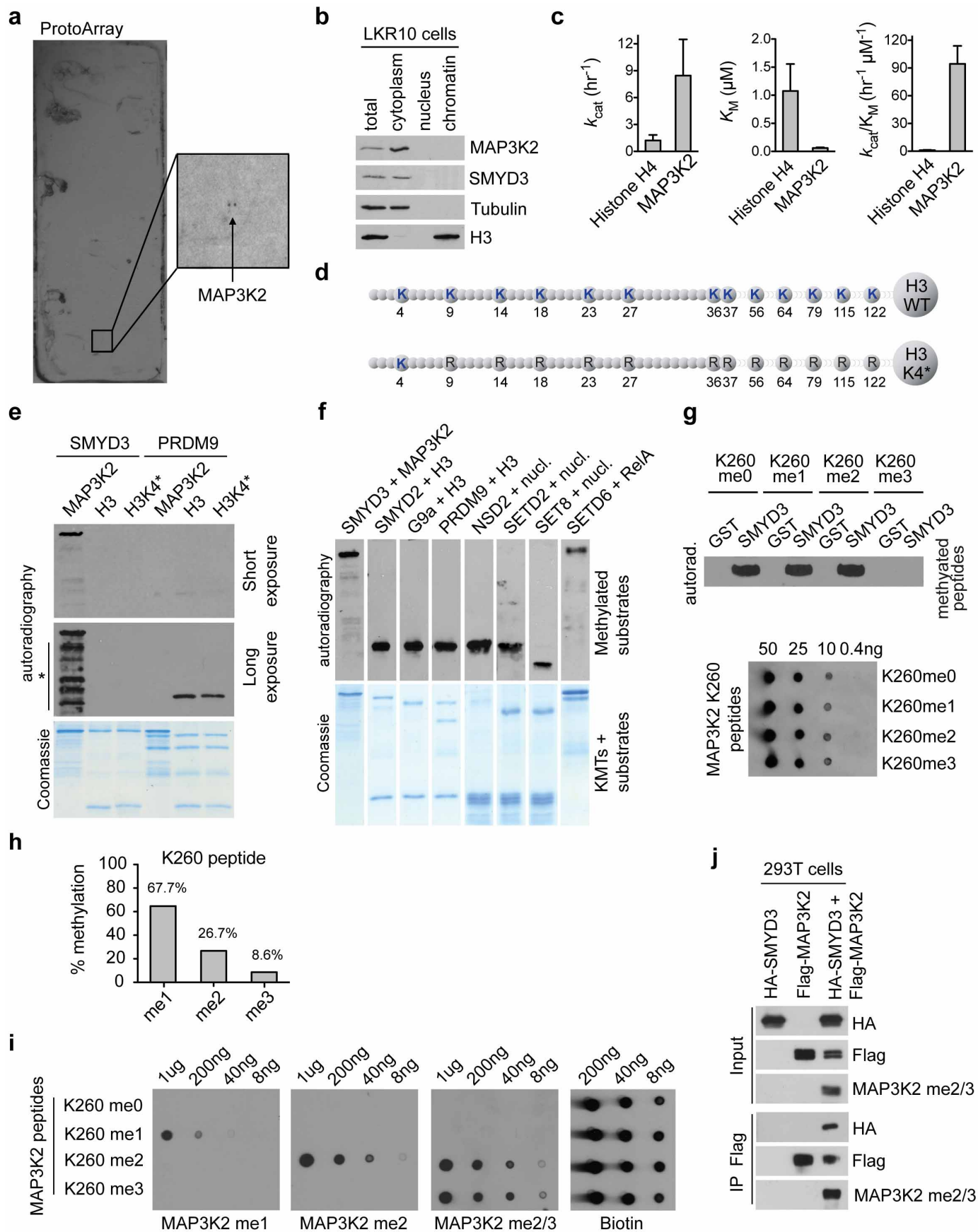
**a**, IHC analysis of SMYD3 reconstitution in the lung (from Fig. 3a).

**b**, Immunofluorescent detection of SMYD3 expression in wild-type and transduced acinar clusters (left panel). Acini (asterisk) transduced with lenti-Cre carrying wild-type SMYD3 but not catalytically inactive

SMYD3(F183A) undergo ADM and form ducts (arrowhead) *ex vivo*.

**c**, Quantification of acinar and ductal clusters after lentiviral infection (each treatment represents four independent biological replicates). Data are represented as mean  $\pm$  s.e.m. \**P* value < 0.05 (two-tailed unpaired Student's *t*-test).

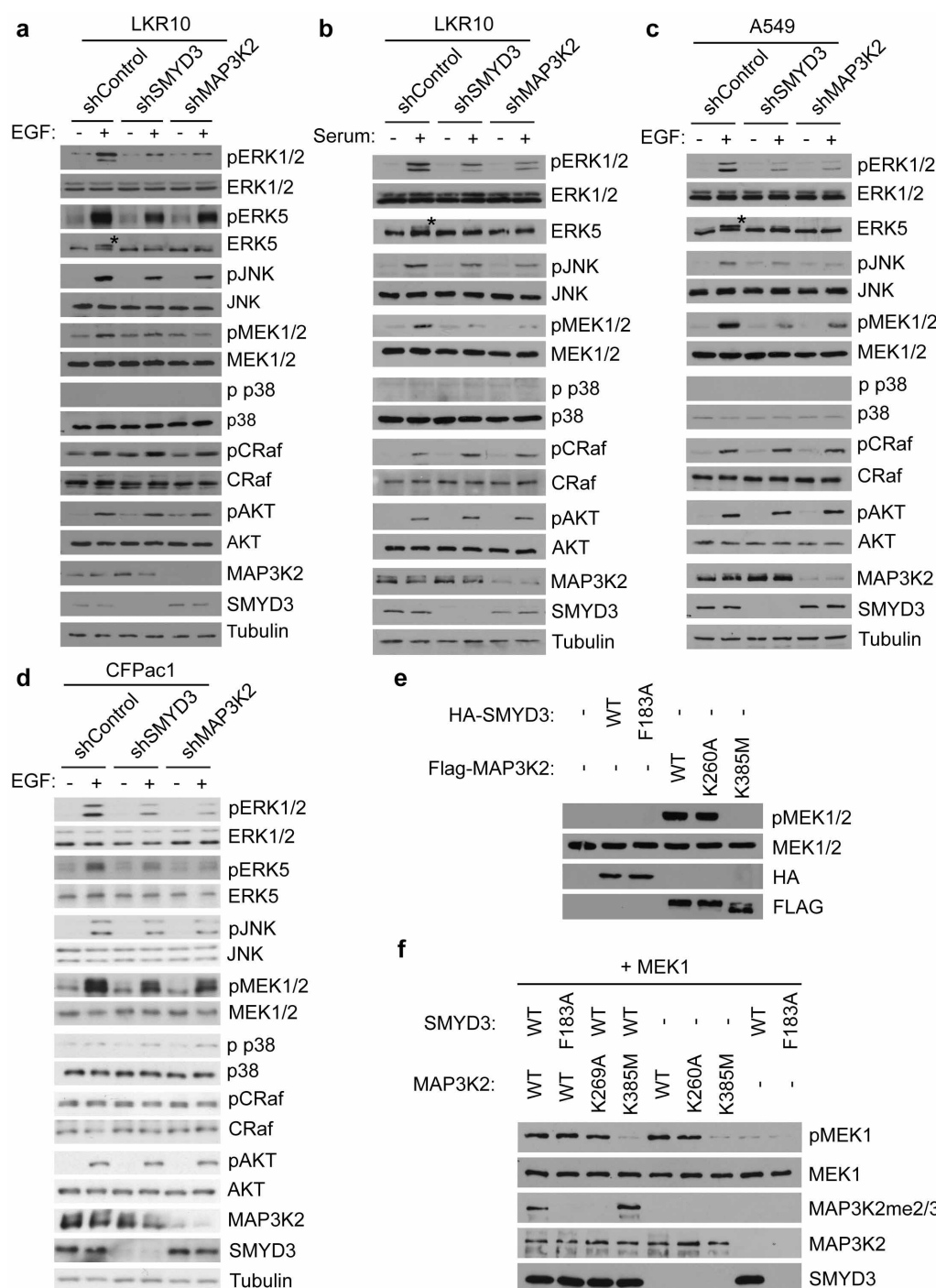






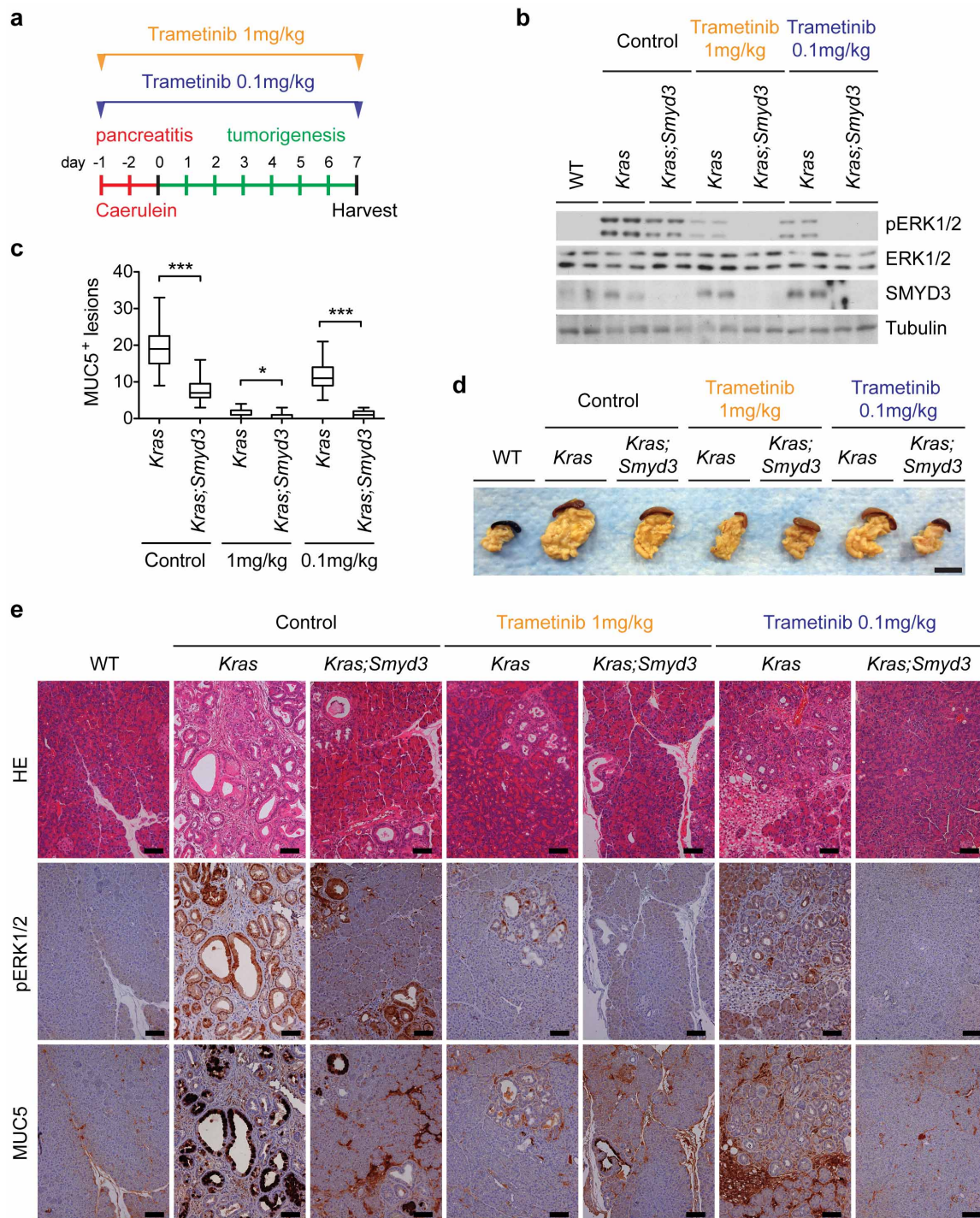
**Extended Data Figure 6 | SMYD3 specifically methylates MAP3K2 at lysine 260 *in vitro*.** **a**, SMYD3 methylates MAP3K2 on protein arrays. Representative image ( $n = 3$  independent experiments) showing a SMYD3 methylation assay on a ProtoArray. The close-up shows the two independent MAP3K2 spots on the array being methylated. **b**, SMYD3 is detected in the cytoplasm and not the nucleus in LKR10 cells. Immunoblot analysis with the indicated antibodies of LKR10 cell lysates biochemically separated into cytoplasmic, nuclear and chromatin fractions (see Methods). **c**, SMYD3 catalytic efficiency is two orders of magnitude greater on MAP3K2 than on H4.  $k_{cat}$ ,  $K_M$ , and  $k_{cat}/K_M$  values of SMYD3 activity on recombinant H4 and MAP3K2 as substrates are shown. **d**, Schematic of the H3K4\* mutant form used in **e**. Note that the only lysine available to be methylated in H3 is present at K4. **e**, *In vitro* methylation assay on full-length recombinant MAP3K2, H3 or H3K4\* with recombinant SMYD3 and PRDM9. Top panels, short and long exposure autoradiograms of the methylation assay. No signal was detected for SMYD3

on H3 and H3K4\* after long exposures. The asterisk and line indicate breakdown products of MAP3K2 that contain K260 and can be detected in this methylation assay upon long exposure. Bottom panel, Coomassie stain of proteins in the reaction. **f**, Positive control of activity for enzymes used in Fig. 3f, g on their known respective substrates (MAP3K2, histone H3, nucleosome or RelA as indicated). **g**, *In vitro* methylation assays using MAP3K2-K260meo, me1, me2 or me3 peptides as SMYD3 substrates. Dot blot is shown as control of peptide's comparable concentration used for the methylation assay. **h**, Mass spectrometry analysis of SMYD3 methylation activity on unmodified MAP3K2-K260 peptide. **i**, Specificity of the indicated MAP3K2-K260me antibodies in dot blot assays using MAP3K2-K260meo, me1, me2 or me3 peptides. **j**, MAP3K2 is methylated in cells upon SMYD3 overexpression. Immunoblot analysis with the indicated antibodies from 293T cells lysates after Flag immunoprecipitation in cells overexpressing Flag-MAP3K2 and/or HA-SMYD3.



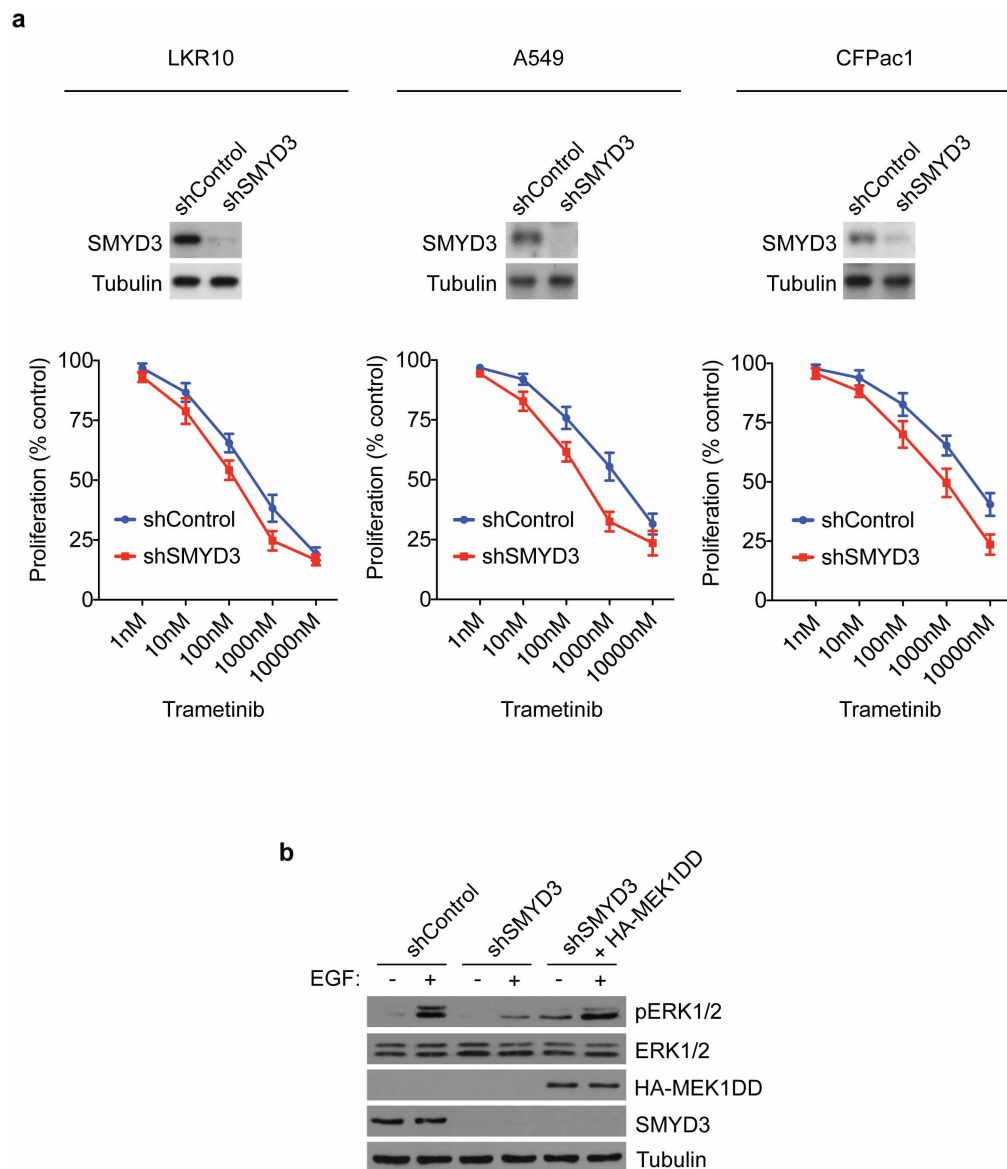
**Extended Data Figure 7 | SMYD3 and MAP3K2 knockdown both impair MAP kinase signalling.** **a–d**, Immunoblot analysis with the indicated antibodies of LKR10 (**a**, **b**), A549 (**c**), and CFPac1 (**d**) lysates. Asterisk indicates a slower migrating ERK5 species that is phosphorylated. Stimulation, 10% serum-complemented media for 15 min (**b**) or EGF for 15 min at  $25 \text{ ng } \mu\text{l}^{-1}$  (**a**, **c**, **d**). Immunoblots are representative of 3 independent biological replicates. **e**, **f**, SMYD3 methylation of MAP3K2 does not alter the intrinsic kinase activity of MAP3K2. **e**, Immunoblot analysis with the indicated antibodies from lysates of 293T cells transfected with control vector, wild-type SMYD3,

catalytically dead SMYD3(F183A), wild-type MAP3K2, MAP3K2(K260A), or kinase dead MAP3K2(K385M). **f**, Methylation of MAP3K2 does not alter its *in vitro* kinase activity. *In vitro* kinase assays were performed with the indicated recombinant versions of MAP3K2 (wild-type, SMYD3-resistant K260A mutant, or kinase dead K385 mutant) pre-methylated with wild-type SMYD3 or as a control, inactive SMYD3, using MEK1 as a substrate. MEK1 phosphorylation was detected by immunoblot analysis with the indicated antibody.



**Extended Data Figure 8 | SMYD3 knockout augments the effects of the MEK1/2 inhibitor trametinib (GSK1120212) *in vivo*.** **a**, Schematic of the caerulein pancreatitis-induced tumorigenesis protocol. Mice were treated with a normal dose of trametinib (1 mg per kg intraperitoneally daily) or a low dose (0.1 mg per kg intraperitoneally daily) or vehicle control. **b**, Immunoblot analysis with indicated antibodies of two independent pancreas biopsies per treatment group. **c**, Quantification of MUC5-positive lesions in

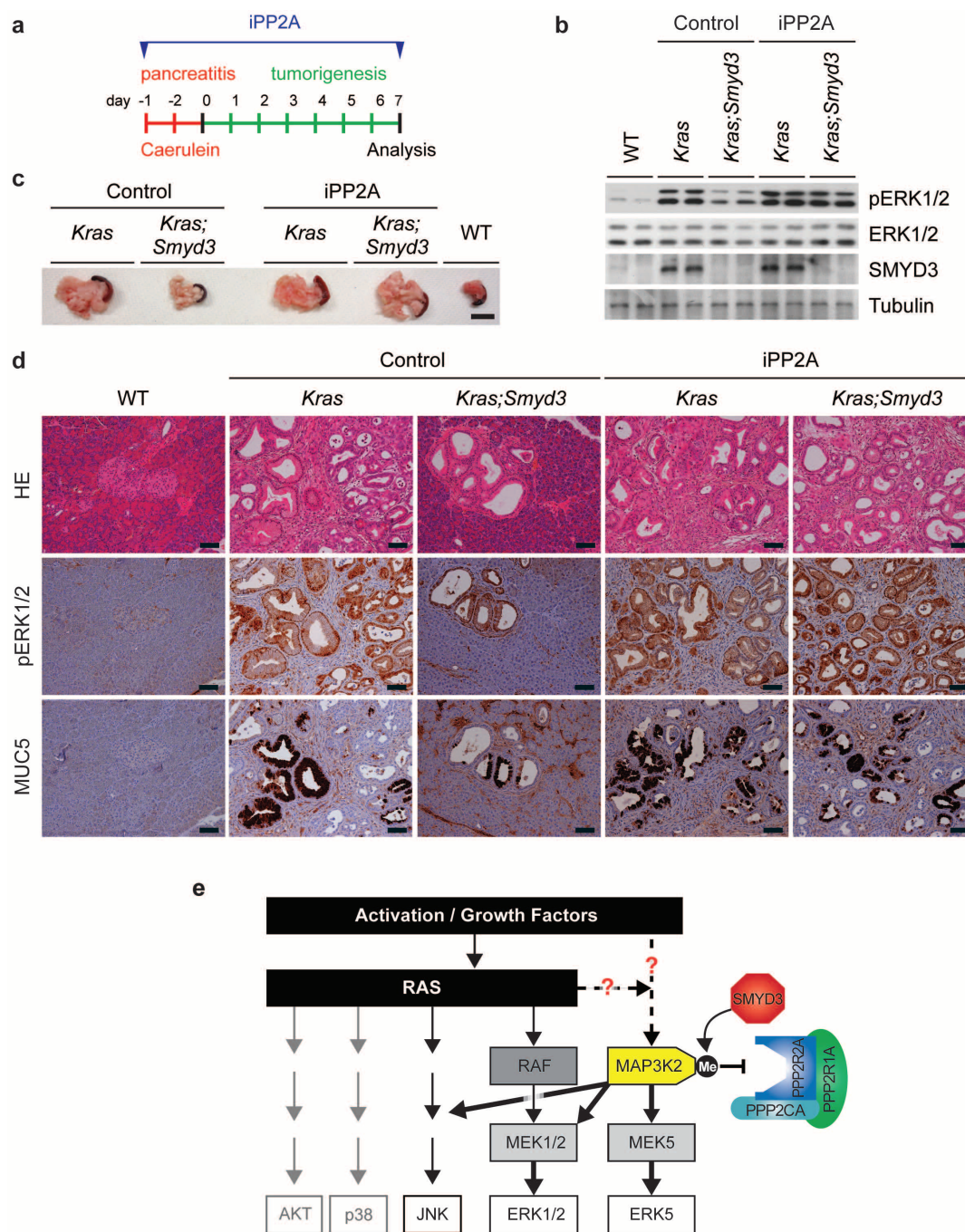
caerulein-treated pancreata from *Kras* and *Kras;Smyd3* mutant mice treated with trametinib or vehicle control ( $n = 5$ , each treatment). \* $P$  value  $< 0.05$ ; \*\*\* $P$  value  $< 0.001$  (two-tailed unpaired Student's  $t$ -test). Data are represented as mean  $\pm$  s.e.m. **d**, Representative macroscopic pictures of pancreata from each treatment group. Scale bar, 1 cm. **e**, Representative serial HE staining and IHC for pERK1/2, a marker of Ras activity, and MUC5, a marker of PanIN lesions. All scale bars, 50  $\mu$ m.



**Extended Data Figure 9 | SMYD3 depletion augments the effects of the MEK1/2 inhibitor trametinib (GSK1120212) in Ras-driven cancer cells.** **a–c**, Relative cell proliferation rates (bottom panel) of murine LAC cell line LKR10 (**a**), human LAC cell line A549 (**b**), or human PDAC cell line CFPac1 (**c**) with or without SMYD3 depletion by stable shRNA (SMYD3 proteins levels are shown in top panel) in response to the indicated doses of trametinib. Experiments shown represent an average of 3 independent experiments

performed in triplicates for each cancer line. Values represent the number of cells relative to control shRNA cells without treatment at 48 h. **b**, Constitutively active MEK1 (MEK1-DD) increases EGF-mediated ERK1/2 activation in SMYD3 depleted-cells. Immunoblot analysis with the indicated antibodies using lysates from A549 cells stably expressing shControl or shSMYD3 and transfected with HA-MEK1-DD. Stimulation: EGF treatment for 15 min at  $25 \text{ ng } \mu\text{l}^{-1}$ .





**Extended Data Figure 10 | Treatment with the PP2A inhibitor cantharidin phenocopies SMYD3 function *in vivo*.** **a**, Schematic of the caerulein pancreatitis-induced tumorigenesis protocol. Mice were treated with the PP2A inhibitor cantharidin (iPP2A, 0.15 mg kg<sup>-1</sup> intraperitoneally twice a day) or vehicle control. **b**, Immunoblot analysis with indicated antibodies on two independent pancreas biopsies per treatment group. **c**, Macroscopic pictures of WT and *Kras;Smyd3* mutant pancreata. Note that treatment with the PP2A inhibitor leads to the development of enlarged, 'hard' pancreata characteristic of tumorigenic development even in *Kras;Smyd3* mutant mice. Scale bar, 1 cm. **d**, Representative serial haematoxylin and eosin (HE) staining and IHC for pERK1/2, a marker of Ras activity, and MUC5, a marker of PanIN lesions. All scale bars, 50  $\mu$ m. **e**, Summary model for SMYD3 regulation of MAP kinase signalling after MAP3K2 methylation. Oncogenic Ras activates several kinase

cascades that play important roles in pancreas and lung cancer development, including four major MAPK pathways (ERK1/2, ERK5, JNK, and p38) as well as AKT signalling. SMYD3 is frequently overexpressed in pancreatic and lung cancers, two cancer types that are commonly driven by oncogenic Ras signalling. Overexpression of SMYD3 and the resulting methylation of MAP3K2 at K260 potentiate activation of kinases like ERK1/2 and ERK5 in response to stimuli like oncogenic Ras. We postulate a mechanism in which the PP2A complex is unable to bind methylated MAP3K2, which decreases the ability of this enzyme to terminate activating phosphorylation events on MAP3K2 and/or MAP3K2 downstream targets. Under conditions with excessive SMYD3 protein, the physiological relationship between PP2A and MAP3K2 is disrupted and results in an increased pathological MAP3K2 signalling, which cooperates with Ras to promote tumorigenesis.

# CTP synthase 1 deficiency in humans reveals its central role in lymphocyte proliferation

Emmanuel Martin<sup>1,2</sup>, Noé Palmic<sup>1,2</sup>, Sylvia Sanquer<sup>3</sup>, Christelle Lenoir<sup>1,2</sup>, Fabian Hauck<sup>1,2</sup>, Cédric Mongellaz<sup>4</sup>, Sylvie Fabrega<sup>2,5</sup>, Patrick Nitschké<sup>2,6</sup>, Mauro Degli Esposti<sup>7,8</sup>, Jeremy Schwartzentruber<sup>9</sup>, Naomi Taylor<sup>4</sup>, Jacek Majewski<sup>9</sup>, Nada Jabado<sup>9,10</sup>, Robert F. Wynn<sup>7</sup>, Capucine Picard<sup>2,11,12</sup>, Alain Fischer<sup>1,2,13,14</sup>, Peter D. Arkwright<sup>7\*</sup> & Sylvain Latour<sup>1,2,3\*</sup>

Lymphocyte functions triggered by antigen recognition and co-stimulation signals are associated with a rapid and intense cell division, and hence with metabolism adaptation<sup>1</sup>. The nucleotide cytidine 5' triphosphate (CTP) is a precursor required for the metabolism of DNA, RNA and phospholipids<sup>2–4</sup>. CTP originates from two sources: a salvage pathway and a *de novo* synthesis pathway that depends on two enzymes, the CTP synthases (or synthetases) 1 and 2 (CTPS1 with CTPS2); the respective roles of these two enzymes are not known<sup>5–7</sup>. CTP synthase activity is a potentially important step for DNA synthesis in lymphocytes<sup>8,9</sup>. Here we report the identification of a loss-of-function homozygous mutation (rs145092287) in *CTPS1* in humans that causes a novel and life-threatening immunodeficiency, characterized by an impaired capacity of activated T and B cells to proliferate in response to antigen receptor-mediated activation. In contrast, proximal and distal T-cell receptor (TCR) signalling events and responses were only weakly affected by the absence of CTPS1. Activated CTPS1-deficient cells had decreased levels of CTP. Normal T-cell proliferation was restored in CTPS1-deficient cells by expressing wild-type *CTPS1* or by addition of exogenous CTP or its nucleoside precursor, cytidine. CTPS1 expression was found to be low in resting T cells, but rapidly upregulated following TCR activation. These results highlight a key and specific role of CTPS1 in the immune system by its capacity to sustain the proliferation of activated lymphocytes during the immune response. CTPS1 may therefore represent a therapeutic target of immunosuppressive drugs that could specifically dampen lymphocyte activation.

We initially studied two unrelated families (family 1 and 2) originating from the northwest region of England, whose four children suffered from severe and recurrent Epstein–Barr virus (EBV) infection, in whom known primary immunodeficiencies have been excluded<sup>10</sup> (Fig. 1a and Table 1). Four additional patients (family 3 to 5) originating from the same geographical area were identified thereafter (Methods). All patients had early onset of severe chronic viral infections, mostly caused by herpes viruses, including EBV and varicella zoster virus (VZV) and also suffered from recurrent encapsulated bacterial infections, a spectrum of infections typical of a combined deficiency of adaptive immunity (CID)<sup>11</sup> (Table 1 and data not shown). Two patients (P4 and P5) had EBV-driven B-cell non-Hodgkin lymphoma. Overall, the clinical phenotype is severe with 3 patients having died. Six of 8 patients have undergone haematopoietic stem cell transplantation. Of note, none of the patients had extra-haematopoietic manifestations (Table 1).

Immunological investigations showed that most of patients had variable lymphopenia which was exacerbated during infection episodes with

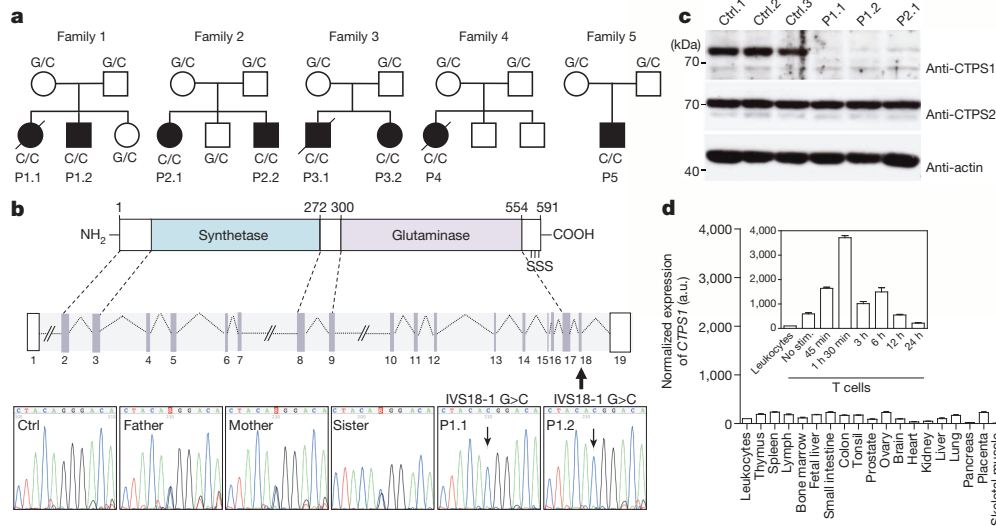
inverted CD4:CD8 T-cell ratio, whereas other blood cell counts were usually normal (Extended Data Table 1 and data not shown). Their immunoglobulin levels were normal or elevated with increased IgG but low IgG2 levels with low antibody titres to *Streptococcus pneumoniae*. Further analyses were performed in patient P1.2 showing naive CD4<sup>+</sup> T-cell lymphopenia, increased numbers of effector memory T cells, low numbers of memory CD27<sup>+</sup> B cells, a complete absence of both invariant T cell populations (CD3<sup>+</sup>Vα24<sup>+</sup>Vβ11<sup>+</sup>) iNKT and (CD3<sup>+</sup>CD161<sup>Hi</sup>Vα7.2<sup>+</sup>) MAIT cells, as well as an impaired PHA- and antigen-induced proliferation of peripheral blood mononuclear cells (PBMCs) (Extended Data Table 2).

To identify the gene defect underlying the immunodeficiency in these patients, we performed whole-exome sequencing (WES) in three patients (P1.1, P1.2 and P2.1). Intersection of the genetic variations found in the three patients pointed to a unique common homozygous G to C mutation in the *CTPS1* gene encoding the CTP synthase 1 at position 41475832 in chromosome 1 with an assigned rsID (rs145092287) in the dbSNP database (Fig. 1b and Extended Data Fig. 1a, b). *CTPS1* encodes a 67-kDa protein containing a CTP synthetase domain and a glutamine amide transfer domain promoting the formation of CTP from UTP and glutamine<sup>12</sup>. The identified mutation affects a splice donor site at the junction of intron 17–18 and exon 18 (IVS18-1 G to C) leading to the expression of an abnormal transcript lacking exon 18 (Extended Data Fig. 1b, c). This splice mutation was found to be deleterious because CTPS1 protein expression could not be detected in lysates of EBV-transformed B cells and T-cell blasts from patients (Figs 1c and 2c and Extended Data Fig. 2). In contrast, CTPS2 was expressed normally in patient cell lysates. In the five affected families, all patients were homozygous for the IVS18-1 G to C mutation and all parents and tested healthy siblings were heterozygous carriers (Fig. 1a, b and data not shown). Sequencing of a cohort of 752 healthy individuals from the northwest of England gave an estimated frequency of homozygosity of 1:560,000. This represents more than a tenfold increase compared to the frequency estimated from available exome databases. WES data and analysis of polymorphic microsatellite markers in all patients revealed a common region of homozygosity of 1.1 Mb surrounding the IVS18-1 G to C mutation (Supplementary Information). All these data were indicative of a founder effect. These observations led us to conclude that the immunodeficiency resulting from the *CTPS1* mutation in these patients could be primarily associated with a T-cell immunodeficiency.

We next examined CTPS1 expression in normal tissues. *CTPS1* mRNA expression was comparable between the different tissues, except for T cells in which *CTPS1* expression was strongly upregulated after cell

<sup>1</sup>Laboratoire Activation Lymphocytaire et Susceptibilité à l'EBV, INSERM UMR 1163, Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>2</sup>Université Paris Descartes Sorbonne Paris Cité, Institut Imagine, Paris 75015, France. <sup>3</sup>Laboratoire de Biochimie Métabolique et Protéomique, Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>4</sup>Hématopoiesis and Immunotherapy, CNRS-UMR 5535, Institut de Génétique Moléculaire de Montpellier, Montpellier 34293, France. <sup>5</sup>Plateforme Vecteurs Viraux et Transfert de Gènes, IFR94, Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>6</sup>Service de Bioinformatique, Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>7</sup>University of Manchester, Royal Manchester Children's Hospital, Manchester M13 0WL, UK. <sup>8</sup>Italian Institute of Technology, Genoa 16163, Italy. <sup>9</sup>McGill University and Genome Québec Innovation Centre, Montréal H3A 0G1, Canada. <sup>10</sup>Department of Pediatrics, McGill University Health Center Research Institute, Montréal H3H 1P3, Canada. <sup>11</sup>Centre d'Etude des Déficiences Immunitaires, Hôpital Necker Enfants-Malades, AP-HP, Paris 75015, France. <sup>12</sup>Laboratoire Génétique Humaine des Maladies Infectieuses, INSERM UMR 1163, Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>13</sup>Unité d'Immunologie et Hématologie Pédiatrique, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Necker Enfants-Malades, Paris 75015, France. <sup>14</sup>Collège de France, Paris 75005, France.

\*These authors contributed equally to this work.



**Figure 1 | Identification of *CTPS1* deficiency in patients with a combined immunodeficiency.** **a**, Pedigrees of the families in which a homozygous IVS18-1 G to C mutation in *CTPS1* was identified. When known, the genotype of each individual is indicated. Black boxes represent affected individuals and diagonal bars indicate deceased individuals. Each patient (P) is identified by a number. **b**, Diagram of the *CTPS1* intron–exon organization and protein domains with the serine phosphorylation sites (S) indicated and the coding exons in grey. DNA electropherograms show the region containing the mutation in *CTPS1* in family 1. The homozygous IVS18-1 G to C mutation is

indicated by an arrow. Ctrl, control. **c**, Immunoblots for CTPS1 and CTPS2 expression in non-stimulated EBV B-cell lines from healthy controls and *CTPS1*-mutated individuals (P1.1, P1.2 and P2.1). Actin serves as a loading control. **d**, *CTPS1* mRNA expression in normal tissues monitored by qRT-PCR in arbitrary units (a.u.). The inset shows the kinetics of *CTPS1* mRNA expression following anti-CD3+CD28 coated beads stimulation. One representative experiment of three. Data represent means of technical triplicates; error bars denote standard deviations.

activation in response to TCR-CD3 and CD28 co-stimulation (Fig. 1d). Interestingly, in lysates from resting non-activated T-cell blasts and T cells from PBMCs, CTPS1 protein was almost undetectable (Fig. 2a–d). In contrast, CTPS2 expression was readily detected. Activation of T cells by anti-CD3 antibody or phorbol 12-myristate 13-acetate (PMA) and ionomycin stimulations induced CTPS1 protein expression, whereas activation with IL-2 and/or IL-15 resulted in only a weak effect. Under the same experimental conditions, CTPS2 expression was also induced but to a lesser extent. In TCR-CD3-stimulated T-cell blasts, CTPS1 protein expression was enhanced from 12 h and persisted for up to 96 h as a consequence of *CTPS1* gene transcription activation (Fig. 1d, inset and Fig. 2b). As expected, no expression of CTPS1 was detected in T-cell blasts from the *CTPS1*-deficient patient (P1.2) contrasting with detection of *CTPS1* mRNA and suggesting protein instability (Fig. 2c and

Extended Data Fig. 1c). These data indicate that T-cell activation through the TCR results in a rapid and sustained CTPS1 protein expression. Of note, in B cells activated by anti-BCR and CpG, IL-4 and CD40L or PMA and ionomycin, CTPS1 was also found to be upregulated (Fig. 2d and Extended Data Fig. 3a, b).

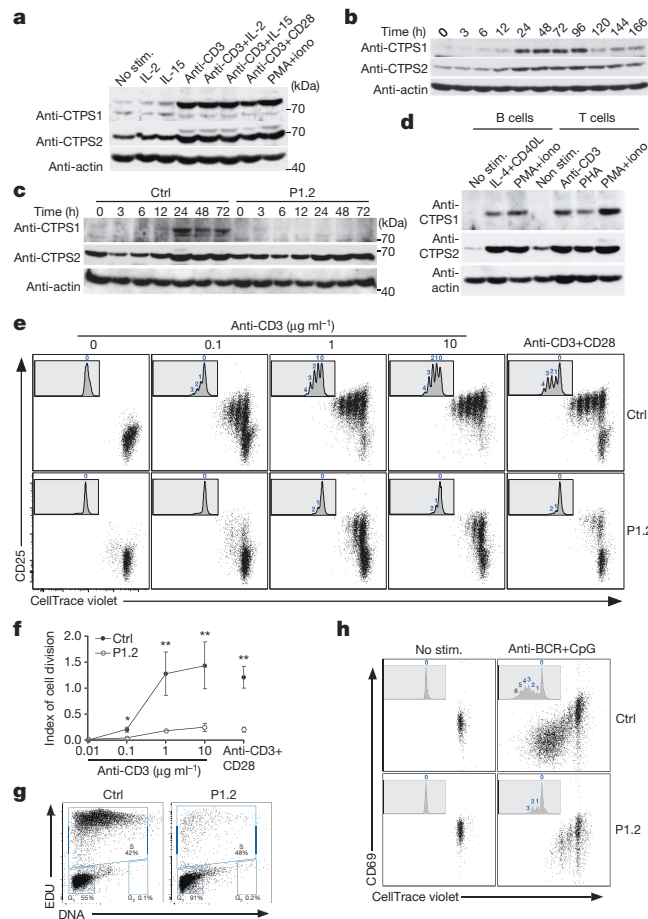
To further characterize the consequences of the *CTPS1* deficiency in T cells, we investigated proximal T-cell activation signals as well as late responses. *CTPS1*-deficient T cells exhibited normal early and late responses with the exception of ERK1/2 phosphorylation and CD25 and CD69 upregulation which were found to be decreased (Extended Data Fig. 4). Basal and activation-induced cell death was also slightly increased (Extended Data Fig. 4g). These data suggest that *CTPS1* deficiency had limited consequences in signalling downstream of TCR-CD3. Because the pool of CTP is potentially a limiting factor for DNA synthesis<sup>8,13</sup>, we

**Table 1 | Clinical features of patients**

Patient	Age at first symptoms	Viral infections			Bacterial infections	Extra- haematopoietic manifestations	Outcome (age in years)
		EBV	VZV	Others			
P1.1	1 y.	SIM, chronic viraemia	No	CMV, Novovirus, Rotavirus (gut)	<i>H. influenzae</i> (RTI)	No	HSCT (8 y.) died (GVHD) (8 y.)
P1.2	1 m.	SIM	No	Parainfluenzae I (RTI)	No	No	Alive (9 y.)
P2.1	5 y.	LPD (CNS)	Yes	Adenovirus, HHV-6, Novovirus (gut)	<i>H. influenzae</i> (RTI)	No	HSCT (9 y.) a.w. (19 y.)
P2.2	2 y.	Chronic viraemia	No	No	<i>S. pneumoniae</i> , <i>H. influenzae</i> (RTI)	No	HSCT (7 y.) a.w. (14 y.)
P3.1	1 y.	SIM	Yes (gastritis, pneumonitis)	No	<i>S. pneumoniae</i> (sepsis, meningitis)	No	Died (disseminated VZV) (4 y.)
P3.2	3 m.	SIM, chronic viraemia	Yes	HHV-6	No	No	HSCT (8 y.) a.w. (14 y.)
P4	Birth	LPD (CNS)	Yes	CMV, Adenovirus, Rotavirus (gut)	No	No	HSCT (6 y.) died (LPD) (6 y.)
P5	3 m.	LPD (CNS), chronic viraemia	No	Novovirus (gut), Parainfluenzae III, Adenovirus, Rhinovirus (RTI)	<i>N. meningitis B</i> (meningitis)	No	HSCT (1 y.) alive (2 y.)

Abbreviations are: y., year; m., month; SIM, severe infectious mononucleosis; CNS, central nervous system; EBV, Epstein-Barr virus; VZV, varicella zona virus; HHV-6, human herpes virus 6; LPD, lymphoproliferative disease; RTI, respiratory tract infection; CMV, cytomegalovirus; HSCT, haematopoietic stem cell transplantation; GVHD, graft versus host disease; a.w., alive and well.





**Figure 2 | Induction of CTPS1 expression during T-cell activation and defective proliferation of activated CTPS1-deficient T cells.**

**a–d**, Immunoblots for CTPS1 and CTPS2 expression in control T cells (from a healthy donor) treated with various stimuli (**a**) or stimulated with anti-CD3 for different periods of time (**b**), in control (Ctrl) or CTPS1-deficient cells from patient P1.2 stimulated with anti-CD3 for different periods of time (**c**) and in B and T cells sorted from normal PBMCs treated with indicated stimuli (**d**). Actin served as a loading control. **e**, Representative dot plots showing cell divisions by dilution of the violet dye and expression of CD25 of control (Ctrl) or CTPS1-deficient T cells (patient P1.2) stimulated with incremental doses of the anti-CD3 antibody or anti-CD3+CD28 coated beads. Inserts with histograms showing the violet dye dilution with the number cell divisions indicated at the top of each peak. Data from one of four independent experiments. **f**, Mean of index values of cell division of control T cells (Ctrl) or CTPS1-deficient cells (P1.2) ( $n = 4$ ). Unpaired Student's  $t$ -tests and  $**P < 0.01$ . **g**, Representative dot plots of cell cycle progression of control (Ctrl) and CTPS1-deficient T-cells (patient P1.2) stimulated with anti-CD3 antibody. The percentages of cells in each stage are indicated. Data from one of two independent experiments. **h**, Proliferation of control (Ctrl) or CTPS1-deficient CD19<sup>+</sup> B cells from PBMCs of healthy donor and patient P1.2. Cells were stimulated with anti-BCR plus CpG during 5 days. The proliferation was analysed similarly as in (e). Representative data from one of two independent experiments.

carefully analysed proliferation of CTPS1-deficient T cells. In response to activation by antigens, anti-CD3 antibody or co-stimulation by anti-CD3 and anti-CD28 antibodies, CTPS1-deficient cells from three patients (P1.1, P1.2 and P2.2) failed to sustain proliferative responses as measured by [<sup>3</sup>H]thymidine uptake and CFSE or violet cell tracer dye dilution (resulting in a weak index of cell proliferation) (Fig. 2e, f and Extended Data Table 1 and Extended Data Figs 5 and 6). Uptakes of [<sup>3</sup>H]uridine and [<sup>3</sup>H]cytidine were also found to be impaired in activated CTPS1-deficient T cells. This suggests that both RNA and DNA synthesis were affected (Extended Data Fig. 6). Defective proliferation of CTPS1-deficient T cells was associated with a lack of cell cycle progression as a majority

of cells were arrested in the G1 phase (Fig. 2g). Proliferation of CTPS1-deficient B cells activated by anti-BCR and CpG was also found to be defective, whereas that of IL-2-activated natural killer cells seemed to be less affected (Fig. 2h and Extended Data Fig. 5b).

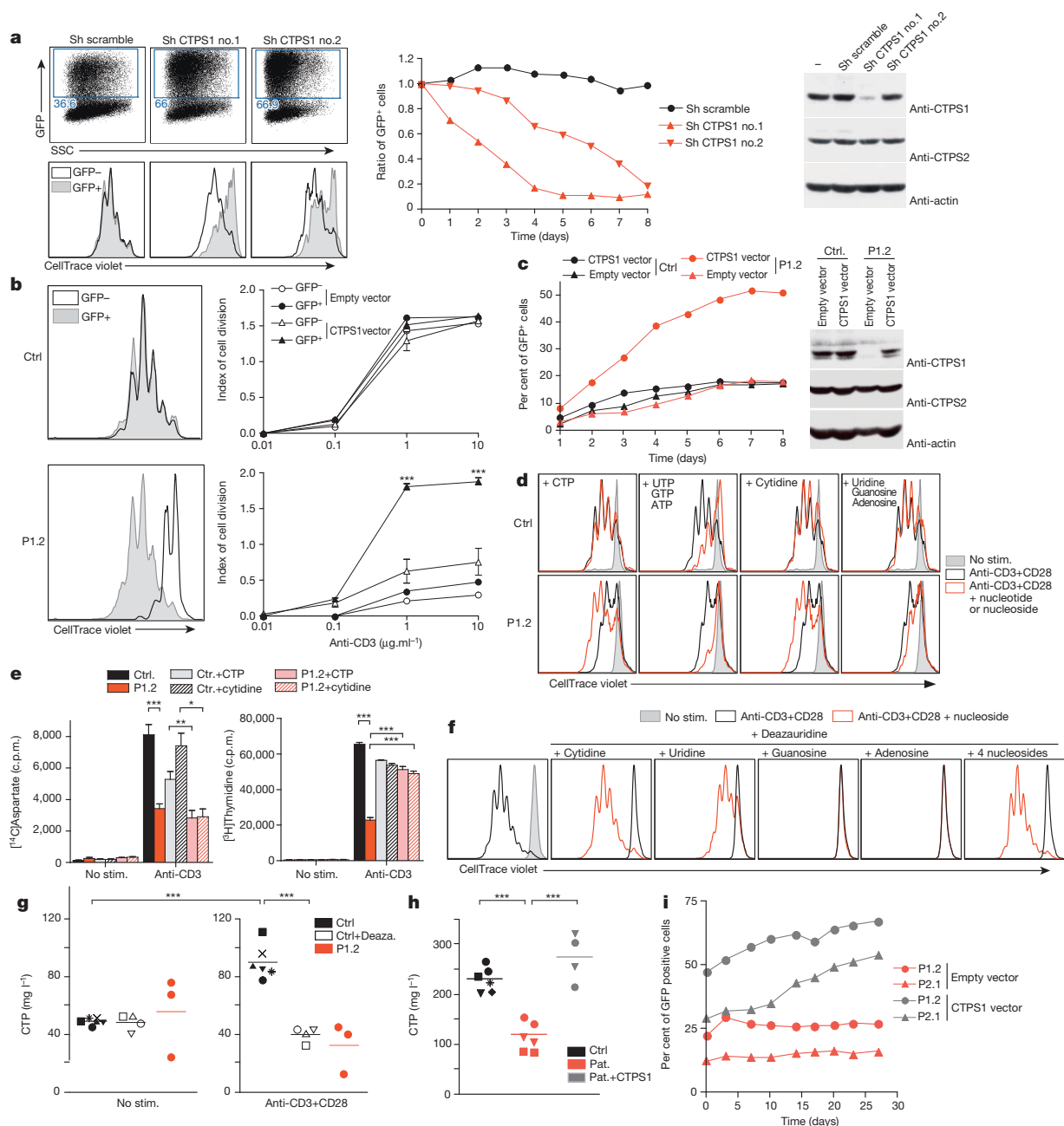
Downregulation of CTPS1 expression in control T cells, by lentiviral transduction of two distinct short hairpin RNA (shRNA) together with a GFP reporter gene, resulted in a specific decrease in the CD3-mediated proliferation of GFP-positive cells (Fig. 3a). No changes in proliferation were detected in non-targeted GFP-negative cells or in cells targeted with a scramble shRNA. The diminished proliferation resulting from the inhibition of CTPS1 expression led to a selective cell growth disadvantage with a decreased number of GFP targeted cells over time (middle panel). A similar decrease in proliferation rate was also observed in the Jurkat T-cell line in which CTPS1 expression was downregulated (Extended Data Fig. 7).

Together, these results indicate that CTPS1 deficiency causes a defect in T-cell proliferation in response to TCR–CD3 activation. To formally prove the causal relationship between CTPS1 deficiency and defective T-cell proliferation, we carried out reconstitution experiments with wild-type CTPS1 or by direct addition of CTP or its cytidine precursor that acts on CTP levels through the salvage pathway. Expression of ectopic CTPS1 in CTPS1-deficient T cells fully restored proliferation upon CD3 stimulation (Fig. 3b) and enabled cells to expand selectively as shown by the accumulation of GFP-positive cells expressing CTPS1 (Fig. 3c, left panels). No such effect was detected in CTPS1-deficient cells transduced with an empty vector or in control cells transduced with the CTPS1-containing vector.

Proliferation and CD25 expression of CTPS1-deficient cells also recovered to a normal level by addition of CTP or cytidine (Fig. 3d and data not shown). In contrast, addition of a mix of UTP, GTP and ATP or uracil, guanosine and adenosine did not result in increased proliferation of CTPS1-deficient cells. To determine the influence of the CTPS1 defect on the *de novo* pyrimidine synthesis pathway, we measured the incorporation of carbon from [<sup>14</sup>C]aspartate into nucleic acids of activated CTPS1-deficient T cells, which is a specific assay for *de novo* pyrimidine synthesis<sup>14</sup> (Fig. 3e and Extended Data Fig. 6b, c). Incorporation of [<sup>3</sup>H]thymidine and [<sup>3</sup>H]uridine was analysed in parallel as control of the global RNA and DNA synthesis. TCR–CD3 activation-mediated incorporation of [<sup>14</sup>C]aspartate, [<sup>3</sup>H]thymidine and [<sup>3</sup>H]uridine was significantly decreased in CTPS1-deficient T cells. Addition of exogenous CTP or cytidine that bypassed the *de novo* synthesis pathway restored incorporation of [<sup>3</sup>H]thymidine but not of [<sup>14</sup>C]aspartate in CTPS1-deficient cells, thus demonstrating that the *de novo* CTP synthesis pathway is impaired in the absence of CTPS1. Deazauridine, an analogue of UTP and a known inhibitor of CTP synthetase activity<sup>15</sup> completely blocked T-cell proliferation of control cells in response to CD3 activation without affecting proximal TCR–CD3-mediated responses, similar to results observed in CTPS1-deficient cells (Fig. 3f and data not shown). As expected, inhibition of T-cell proliferation by deazauridine was fully reverted by addition of cytidine and partially by uridine, but not by adenosine or guanosine. Analysis of nucleotides pools in activated CTPS1-deficient T-cell blasts and CTPS1-deficient B/EBV cell lines revealed decreased levels of CTP, as also observed in activated normal cells treated with deazauridine (Fig. 3g, h). Defective CTPS1 expression or addition of deazauridine also led to reduced pools of ATP, GTP and UTP in activated T cells (Extended Data Fig. 8) suggesting interconnection in the nucleotide pools<sup>16</sup>. In contrast, CTP as well as ATP, GTP and UTP were found to be normal or increased in resting CTPS1-deficient T cells as the salvage pathway is suggested to be predominant in quiescent cells<sup>9,17</sup>. Expression of wild-type CTPS1 in CTPS1-deficient B/EBV cell lines restored levels of CTP comparable to control cells and conferred a selective growth advantage to cells (Fig. 3h, i).

This study reveals a critical role for CTPS1 in promoting the proliferation of T cells following their activation. However, proliferation of B cells was also found to be dependent on CTPS1. This may directly participate in the susceptibility to encapsulated bacterial infections seen





**Figure 3 | CTPS1 is required for proliferation of T cells in response to TCR-CD3 activation.** **a**, Proliferation of T cells in which *CTPS1* expression was silenced with vectors containing shRNA for *CTPS1* (Sh CTPS1 no. 1 or Sh CTPS1 no. 2) or containing a scramble shRNA (Sh scramble) with GFP gene reporter. Representative dot plots of GFP<sup>+</sup> cells corresponding to transduced cells (left upper panels). Representative histograms of violet dye dilution showing the cell divisions after stimulation (left lower panels). Curves showing the ratio of the percentage of GFP<sup>+</sup> cells at different days to the percentage of GFP<sup>+</sup> cells at day 0 in long-term expansions after repeated stimulation (middle panel). Immunoblots for CTPS1 and CTPS2 expression in transduced cells and non-transduced cells (–) (right panels) and actin as loading control. One representative of two experiments. **b**, **c**, Proliferation of control (Ctrl) and CTPS1-deficient T cells (patient P1.2) transduced by empty or wild-type *CTPS1*-containing vector. Representative histograms of violet dye dilution (b, left panels). Means of indexes of cell division after stimulation (b, right panels) from triplicate of one representative of two experiments. Curves showing the percentage of GFP<sup>+</sup> transduced cells after repeated stimulation (c, left panel). Representative data from one of two independent experiments. Immunoblots same as in (a) (c, right panels). **d**, Representative histograms of violet dye dilution showing cell divisions of control (Ctrl) and

CTPS1-deficient T cells (P1.2) incubated with the indicated nucleotides or nucleosides before stimulation. Data from one of three independent experiments. **e**, Incorporation of [<sup>14</sup>C]aspartate, a tracer of the *de novo* pyrimidine nucleotide synthesis and [<sup>3</sup>H]thymidine as a control of proliferation/DNA synthesis. T cells were labelled during stimulation. Means of incorporated radioactivity (c.p.m.) ( $n = 6$ ). **f**, Same as **d** except that control T cells were incubated with deazauridine. Data from one representative of three independent experiments. **g**, Concentration of CTP in control T cells incubated with deazauridine (Ctrl + Deaza.,  $n = 4$ ) or not (Ctrl,  $n = 6$ ) and CTPS1-deficient cells (P1.2,  $n = 3$ ) after stimulation with anti-CD3+CD28 coated beads. Data from three independent experiments. **h**, Concentration of CTP in cell extracts of EBV B-cell lines from healthy controls (Ctrl;  $n = 6$ ), and CTPS1-deficient patients transduced (Pat. + CTPS1,  $n = 4$ ) or not (Pat.,  $n = 6$ ) with wild-type *CTPS1*-containing vector. P1.1 (squares), P1.2 (circles) and P2.1 (triangles). For controls, each symbol corresponds to cells of a different donor. Data from two independent experiments. **i**, Proliferation of CTPS1-deficient EBV B-cell lines (P1.2 and P2.1) transduced by empty or wild-type *CTPS1*-containing vector. Curves showing the percentage of GFP<sup>+</sup>-transduced cells in culture. Unpaired *t*-tests and \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$  (b, e, g, h).

in CTPS1-deficient patients and account for the low titres of *S. pneumoniae* antibodies as it is a T-cell-independent B-cell response. The role of CTPS1 in B cells could be different or/and less important than in T cells. Of note, CTPS1-deficient B cells preserve an intact capacity to expand upon transformation by EBV and patients had normal immunoglobulin levels and/or elevated IgG. Decreased expansion of natural killer cells and low numbers of iNKT and MAIT cells might also contribute to the CTPS1 immunodeficiency as these cells have been proposed to have a role in a broad range of immune responses including anti-EBV immunity<sup>18–21</sup>. The finding that CTPS1-deficiency causes no other significant clinical consequences favours a redundancy with CTPS2 activity in other cell lineages and tissues. Interestingly, pyrimidine pools including CTP have been previously shown to be strongly expanded in PHA-stimulated T cells by *de novo* pathways including increased CTPS activity<sup>8,9</sup>. The induction of CTPS1 expression in activated T cells reported here thus appears as the major determinant of CTP pool increase. Indeed, proliferation of CTPS1-deficient T cells was restored to normal levels by addition of CTP. The exact mechanism(s) by which TCR signalling induces a rapid expression and activation of CTPS1 in T cells remains to be determined, although we showed that the ERK pathway is required, as well as tyrosine phosphorylation signals (Extended Data Fig. 3c). It is interesting to note that T-cell differentiation does not appear to be severely impaired by CTPS1 deficiency, suggesting that CTP pools in thymocytes may originate from the nucleoside salvage pathway and/or the CTPS2 activity<sup>8,22–24</sup>. Notably, CTPS1 activity is critical for the intense cell division induced by antigenic stimulation as exemplified by massive proliferation and expansion of CD8<sup>+</sup> T cells during viral infections<sup>25,26</sup>.

In the absence of CTPS1, we showed that *de novo* pyrimidine synthesis pathway is impaired but not totally abrogated. This residual activity is probably dependent on CTPS2. Recently, the *de novo* pyrimidine synthesis pathway was shown to be dependent on post-transcriptional regulation by mTORC1 and S6 protein (S6K) kinases that activate the first enzymatic steps required for pyrimidine synthesis<sup>14,27,28</sup>. Thus, distinct regulatory mechanisms control *de novo* pyrimidine synthesis. Based on the present study, CTPS1-mediated tuning of CTP synthesis in lymphocytes appears to be a key element in enabling adaptive immune responses. If CTPS1-specific inhibitors can be designed, they would potentially be highly specific immunosuppressive drugs able to inhibit auto- or allogenic-specific T-cell and B-cell responses without additional toxicity given the lymphocyte specificity of the CTPS1-deficiency phenotype. In conclusion, our results provide the first *in vivo* evidence of a role of the *de novo* pyrimidine synthesis pathway as a critical step for proliferation of T and B lymphocytes when activated by antigens.

## METHODS SUMMARY

Informed consent was obtained from donors, patients and families of patients. The study and protocols conform to the 1975 declaration of Helsinki as well as to local legislation and ethical guidelines. See Methods for full experimental procedures. In several experiments, data are expressed as means  $\pm$  standard deviation (s.d.) denoted by the error bars. *P* values were calculated by two-tailed Student's *t*-test.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 June 2013; accepted 22 April 2014.

Published online 28 May 2014.

- MacIver, N. J., Michalek, R. D. & Rathmell, J. C. Metabolic regulation of T lymphocytes. *Annu. Rev. Immunol.* **31**, 259–283 (2013).
- Evans, D. R. & Guy, H. I. Mammalian pyrimidine biosynthesis: fresh insights into an ancient pathway. *J. Biol. Chem.* **279**, 33035–33038 (2004).
- Higgins, M. J., Graves, P. R. & Graves, L. M. Regulation of human cytidine triphosphate synthetase 1 by glycogen synthase kinase 3. *J. Biol. Chem.* **282**, 29493–29503 (2007).
- Ostrand, D. B., O'Brien, D. J., Gorman, J. A. & Carman, G. M. Effect of CTP synthetase regulation by CTP on phospholipid synthesis in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 18992–19001 (1998).
- Kassel, K. M., Au D. R., Higgins, M. J., Hines, M. & Graves, L. M. Regulation of human cytidine triphosphate synthetase 2 by phosphorylation. *J. Biol. Chem.* **285**, 33272–33276 (2010).

- Nadkarni, A. K. *et al.* Differential biochemical regulation of the URA7- and URA8-encoded CTP synthetases from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **270**, 24982–24988 (1995).
- van Kuilenburg, A. B., Meinsma, R., Vreken, P., Waterham, H. R. & van Gennip, A. H. Identification of a cDNA encoding an isoform of human CTP synthetase. *Biochim. Biophys. Acta* **1492**, 548–552 (2000).
- Fairbanks, L. D., Bofill, M., Ruckemann, K. & Simmonds, H. A. Importance of ribonucleotide availability to proliferating T-lymphocytes from healthy humans. Disproportionate expansion of pyrimidine pools and contrasting effects of *de novo* synthesis inhibitors. *J. Biol. Chem.* **270**, 29682–29689 (1995).
- van den Berg, A. A. *et al.* Cytidine triphosphate (CTP) synthetase activity during cell cycle progression in normal and malignant T-lymphocytic cells. *Eur. J. Cancer* **31**, 108–112 (1995).
- Wynn, R. F. *et al.* Treatment of Epstein-Barr-virus-associated primary CNS B cell lymphoma with allogeneic T-cell immunotherapy and stem-cell transplantation. *Lancet Oncol.* **6**, 344–346 (2005).
- Notarangelo, L. D. Functional T cell immunodeficiencies (with T cells present). *Annu. Rev. Immunol.* **31**, 195–225 (2013).
- Kursula, P. *et al.* Structure of the synthetase domain of human CTP synthetase, a target for anticancer therapy. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62**, 613–617 (2006).
- Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1–22 (1994).
- Ben-Sahra, I., Howell, J. J., Asara, J. M. & Manning, B. D. Stimulation of *de novo* pyrimidine synthesis by growth signaling through mTOR and S6K1. *Science* **339**, 1323–1328 (2013).
- McPartland, R. P., Wang, M. C., Bloch, A. & Weinfeld, H. Cytidine 5'-triphosphate synthetase as a target for inhibition by the antitumor agent 3-deazauridine. *Cancer Res.* **34**, 3107–3111 (1974).
- Qiu, Y. *et al.* Mycophenolic acid-induced GTP depletion also affects ATP and pyrimidine synthesis in mitogen-stimulated primary human T-lymphocytes. *Transplantation* **69**, 890–897 (2000).
- van den Berg, A. A. *et al.* The roles of uridine-cytidine kinase and CTP synthetase in the synthesis of CTP in malignant human T-lymphocytic cells. *Leukemia* **8**, 1375–1378 (1994).
- Le Bourhis, L., Mburu, Y. K. & Lantz, O. MAIT cells, surveyors of a new class of antigen: development and functions. *Curr. Opin. Immunol.* **25**, 174–180 (2013).
- Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nature Immunol.* **9**, 503–510 (2008).
- Chung, B. K. *et al.* Innate immune control of EBV-infected B cells by invariant natural killer T cells. *Blood* **122**, 2600–2608 (2013).
- Brennan, P. J., Brigl, M. & Brenner, M. B. Invariant natural killer T cells: an innate activation scheme linked to diverse effector functions. *Nature Rev. Immunol.* **13**, 101–117 (2013).
- Toy, G. *et al.* Requirement for deoxycytidine kinase in T and B lymphocyte development. *Proc. Natl Acad. Sci. USA* **107**, 5551–5556 (2010).
- Marijn, Y. M. *et al.* Studies on the incorporation of precursors into purine and pyrimidine nucleotides via 'de novo' and 'salvage' pathways in normal lymphocytes and lymphoblastic cell-line cells. *Biochim. Biophys. Acta* **1012**, 148–155 (1989).
- Austin, W. R. *et al.* Nucleoside salvage pathway kinases regulate hematopoiesis by linking nucleotide metabolism with replication stress. *J. Exp. Med.* **209**, 2215–2228 (2012).
- Murali-Krishna, K. *et al.* Counting antigen-specific CD8 T cells: a reevaluation of bystander activation during viral infection. *Immunity* **8**, 177–187 (1998).
- Hislop, A. D., Taylor, G. S., Sauce, D. & Rickinson, A. B. Cellular responses to viral infection in humans: lessons from Epstein-Barr virus. *Annu. Rev. Immunol.* **25**, 587–617 (2007).
- Huang, M. & Graves, L. M. *De novo* synthesis of pyrimidine nucleotides; emerging interfaces with signal transduction pathways. *Cell. Mol. Life Sci.* **60**, 321–336 (2003).
- Robitaille, A. M. *et al.* Quantitative phosphoproteomics reveal mTORC1 activates *de novo* pyrimidine synthesis. *Science* **339**, 1320–1323 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The authors thank the patients, their families and the healthy donors for cooperation. We thank S. Rigaud, S. G  rart, C. Syn  ve and R. Rodriguez for help with experiments and P. Revy for discussion. This work was supported by grants from INSERM, ANR (ANR-08-MIEN-012-01, ANR-2010-MIDI-005-02 and ANR-10-IAHU-01), Fondation ARC (France), the European Research Council (ERC-2009-AdG\_20090506 n  FP7-249816) and the Rare Diseases Foundation (France). S. L. is a senior scientist of CNRS (France). E. M. is supported by ANR (France) and Ligue contre le cancer (France). We are also grateful to the UK 1958 Birth Cohort (<http://www2.le.ac.uk/projects/birthcohort>) for providing DNA from 752 individuals born in the northwest of England. Access to these resources was enabled via the 58READIE Project funded by the Wellcome Trust and Medical Research Council (grant numbers WT095219MA and G1001799). A full list of the financial, institutional and personal contributions to the development of the 1958 Birth Cohort Biomedical resource is available at (<http://www2.le.ac.uk/projects/birthcohort>).

**Author Contributions** E.M. performed experiments, analysed the data and participated in the writing of the manuscript. N.P., F.H., C.L., S.F., C.M. and S.S. performed experiments. A.F., S.L., S.S., J.S., C.P., P.N., J.M., N.J., C.M. and N.T. analysed the data. M.D.E., R.F.W. and P.D.A. identified the families and provided and analysed clinical information. S.L. and A.F. co-wrote the manuscript. S.L. designed and supervised the research.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.L. ([sylvain.latour@inserm.fr](mailto:sylvain.latour@inserm.fr)).

## METHODS

**Cohorts of patients.** Beside the four initially identified patients, four additional patients were identified by screening 10 patients (9 families) originating from the northwest of England with severe chronic viral infections, mostly caused by herpes viruses, including EBV and VZV. Furthermore, 24 patients (24 families) originating from different geographical areas with the same phenotype were also tested for all exons of *CTPS1* in order to identify other mutations and none was found to be a carrier of *CTPS1* mutations.

**Exome sequencing and analysis.** Exome capture was performed according to the manufacturer's protocol using the Illumina TruSeq exome enrichment kit and sequencing of 100 bp paired-end reads on an Illumina HiSeq. Approximately 10 Gb of sequence were obtained for each subject such that 90% of the coding bases of the exome defined by the consensus coding sequence (CCDS) project were covered by at least 10 reads. Adaptor sequences and quality trimmed reads were removed using the Fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)) and a custom script was then used to ensure that only read pairs with both mates present were subsequently used. Reads were aligned to hg19 with BWA31, and duplicate reads were marked using Picard (<http://picard.sourceforge.net/>) and excluded from downstream analyses. Single nucleotide variants (SNVs) and short insertions and deletions (indels) were determined using SAMtools (<http://samtools.sourceforge.net/>) pileup and varFilter32 with the base alignment quality (BAQ) adjustment disabled, and they were then quality filtered to require at least 20% of reads supporting the variant call. Variants were annotated using both ANNOVAR33 and custom scripts to identify whether they affected protein coding sequences, and whether they had previously been seen in the dbSNP, the 1000 Genomes data set (1,092 genomes), or in approximately 2,073 exomes previously sequenced at our centre. A variant detected in a patient was considered to be a candidate mutation if it had not been reported or had a frequency below 0.001% in the three databases indicated above. At the time the homozygous G to C mutation in *CTPS1* (at position chr1:41475832) was identified by WES in the three patients (P1.1, P1.2 and P2.1), it was not described as a dbSNP or assigned to a rsID. Afterwards, this mutation has been identified in the NHLBI GO Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>) with the assigned rsID: rs145092287. In the NHLBI GO Exome Sequencing Project, the rs145092287 is present three times in a heterozygous status among 4,300 genomes from an European-American population and not found in the 2,203 genomes of an African-American population. The rs145092287 was not found in a homozygous or heterozygous status in other available genome databases (NCBI, 1000 Genomes project and the 3,519 genomes of our centre). Homozygosity regions around the rs145092287 were determined in the exomes by looking at the homozygous variations. Between the positions chr1:40737516 (rs6677717) and chr1:42008077 (rs 63729761) a succession of 97 homozygous variations (without heterozygous variations) was found to be shared by the three patients.

**Genomic DNA sequencing.** Genomic DNA from peripheral blood cells, EBV-B cell lines and/or fibroblasts of patients, their parents, and other family members was isolated according to standard methods. Genomic DNAs of 752 healthy control subjects born in the northwest of England were obtained from the UK 1958 Birth Cohort (<http://www2.le.ac.uk/projects/birthcohort/>). The estimated frequency of the *CTPS1* mutation in the populations was calculated according to the Hardy-Weinberg law. Oligonucleotide primers flanking the 3' region of intron 17–18 and exon 18 of *CTPS1* were used to amplify genomic DNA: forward 5'-AGAGTTGGTGGTAGGGTGTGTGAC-3' and reverse 5'-CTTGCAATCGCAGTGTGTTATCAC-3'. PCR products were amplified using high fidelity Platinum Taq DNA Polymerase (Invitrogen) according to the manufacturer's recommendations, purified with the QIAquick gel extraction kit (Qiagen) and sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PerkinElmer) according to the manufacturer's recommendations. All collected sequences were analysed using 4peaks software (Version 1.7.2; A. Griekspoor and T. Groothuis, <http://nucleobytes.com/index.php/4peaks>).

**Analysis of microsatellite markers.** Microsatellite markers were genotyped using UniSTS sequences and mapping information available from the NCBI (<http://www.ncbi.nlm.nih.gov>). Genomic DNA from patients was used as templates to amplify by PCR with specific fluorescent labelled oligonucleotides, the polymorphic repeats corresponding to the microsatellite markers. PCR products were evaluated using an ABI 3100 DNA Fragment Analyzer (Applied Biosystems), and data were analysed using Genescan and Genotyper software (Applied Biosystems).

**Gene expression analysis.** Total RNA was isolated from EBV-B cell and activated T-cell blasts of P1.2, P2.1 patients and control donors using the RNeasy Mini kit (Qiagen). The samples were depleted of genomic DNA and reverse transcription was performed using Superscript II First Strand Synthesis System (Invitrogen). cDNAs were used as a template to perform PCR amplifications of exon 15 to exon 19 of *CTPS1* or exon 4 of actin with the following primers using standard protocols: *CTPS1* forward primer: 5'-GAGAGGCACGCCACCGATTTG-3', *CTPS1* reverse primer: 5'-GCCAGTACACGTGATGGGACATGC-3', actin forward primer:

5'-CTCCTTAATGTACACGACGAT-3'; actin reverse primer: 5'-CTCCTTAATGTCACGACGAT-3'. PCR to amplify full length *CTPS1* cDNAs from control and patients cells were also performed using the following primers, forward primer: 5'-AGCTCTGTGCTGACGGGAGGAT-3' (exon1); reverse primer: 5'-GCCA GTACACGTGATGGGACATGC-3' (exon 19). PCR products were verified by sequencing revealing the expression of an abnormal *CTPS1* transcript lacking exon 18 in patients' cells. Multiple tissue cDNA panels from Ozyme (Human MTC panel I, II and Human Immune Sytem MTC panel) were analysed for *CTPS1* and *CTPS2* gene expression by qRT-PCR. Gene expression assays were performed with Assays-on-Demand probe and primer combinations (*CTPS1*, Hs01041858; *CTPS2*, Hs00219845; *GAPDH*, Hs027558991) from Applied Biosystem labelled with 6-carboxy-fluorescein (FAM) dye, and universal reaction mixture. Real time quantitative PCRs for *GAPDH*, *CTPS1* and *CTPS2* were performed in triplicate using a LightCycler VIIA7 System (Roche). Expression levels were determined by relative quantification using the comparative threshold cycle method  $2^{-\Delta\Delta Ct}$  in which  $\Delta\Delta Ct$  is determined as followed:  $(Ct_{\text{target gene}} - Ct_{\text{reference gene}})_{\text{target tissue}} - (Ct_{\text{target gene}} - Ct_{\text{reference gene}})_{\text{calibrator tissue}}$ . The results shown in arbitrary units (a.u.) have been normalized to *GAPDH* gene expression and are presented as the relative change in gene expression normalized against the calibrator sample corresponding to leukocyte tissue.

**Cell culture and stimulation.** Peripheral blood mononuclear cells (PBMCs) collected from patients and healthy donors were isolated by Ficoll-Paque density gradient (Lymphoprep, Proteogenix) from blood samples using standard procedures. Expansion of T-cell blasts were obtained by incubating PBMCs for 72 h with phytohaemagglutinin (PHA) ( $2.5 \mu\text{g ml}^{-1}$ , Sigma-Aldrich) in RPMI 1640 GlutaMax medium (Invitrogen) supplemented with 5% human male AB blood group serum (BioWest), penicillin ( $100 \text{ U ml}^{-1}$ ) and streptomycin ( $100 \mu\text{g ml}^{-1}$ ). After 3 days, dead cells were removed by Ficoll-Paque density gradient and blasts were maintained in culture with IL-2 ( $100 \text{ U ml}^{-1}$ ). For proliferation and cell cycle analyses, blasts were washed and cultured without IL-2 for 72 h to synchronize the cells. Blasts or PBMCs were then cultured during 4 to 6 days in complete medium alone or in the presence of 0.1, 1 or  $10 \mu\text{g ml}^{-1}$  coated anti-CD3 antibody (clone OKT3, eBiosciences), anti-CD3/CD28-coated beads (Invitrogen), PHA ( $2.5 \mu\text{g ml}^{-1}$ ,  $10^{-5} \text{ M}$  ionomycin (Sigma-Aldrich) plus  $10^{-7} \text{ M}$  phorbol 12-myristate 13-acetate (PMA, Sigma-Aldrich), Candidin ( $5 \mu\text{g ml}^{-1}$ , Bio-Rad), tetanus toxoid (1 to 8,000 dilution, Statens Serum Institute) or tuberculosis antigen ( $50 \mu\text{g ml}^{-1}$ , Statens Serum Institute). Proliferation and cell cycle were analysed at the indicated time points. Then  $40 \mu\text{M}$  of 3-Deazauridine (DAZ, Sigma-Aldrich) was added for 12 h before the stimulation. In complementation experiments, blasts were incubated with  $100 \mu\text{M}$  of CTP, UTP, GTP or ATP (New England Biolabs) separately or in combination, or with  $200 \mu\text{M}$  of cytidine, uridine, guanosine or adenosine (Sigma-Aldrich) separately or in combination 12 h before the stimulation. For dosage of nucleotides, blasts were deprived of IL-2 for 72 h before stimulation or not with anti-CD3/CD28-coated beads for 48 h and cell lysates were prepared. Jurkat cells<sup>29</sup>, 293-T cells and B EBV cell lines from patients were cultured in RPMI 1640 GlutaMax medium supplemented with 10% heat-inactivated fetal calf serum (Gibco), penicillin ( $100 \text{ U ml}^{-1}$ ) and streptomycin ( $100 \mu\text{g ml}^{-1}$ ). Cells were free of mycoplasma contamination.

**Proliferation and cell cycle assays.** Cell proliferation was monitored by labelling cells with the cell trace violet dye (CellTrace violet proliferation kit, Invitrogen) or CFSE ( $5 \mu\text{M}$ , Invitrogen) before stimulation, according to the manufacturer's instructions. After 4 or 6 days of culture, cells were collected and CellTrace violet or CFSE fluorescence dilution was assessed by flow cytometry. The division index of proliferation was calculated using FlowJo software (Tree Star) and corresponds to the average number of cell divisions per cell including the undivided peak. T-cell responses within total PBMCs were also measured by [ $^3\text{H}$ ]thymidine incorporation after 6 days of stimulation. A total of  $0.074 \text{ MBq ml}^{-1}$  of [ $^3\text{H}$ ]thymidine was added during the last 18 h of stimulation. Cell proliferation was determined by c.p.m. of [ $^3\text{H}$ ]thymidine incorporated in cells that were counted with TopCount NXT beta counter (PerkinElmer). Cell cycle analysis was determined by measuring the incorporation of the nucleoside analogue 5-ethynyl-2-deoxyuridine (EdU) into newly synthesized DNA, according to the manufacturer's instructions (Click-iT EdU, Invitrogen) after 48 h of anti-CD3 stimulation. EdU incorporation in cells was measured following conjugation of EdU to azide-modified Alexa Fluor 647 dye. Cells were analysed by flow cytometry with a FACS-Canto II flow cytometry system (BD Biosciences).

**Nucleic acids and de novo pyrimidine synthesis assays.** PBMCs were stimulated in the presence of  $1 \mu\text{g ml}^{-1}$  coated anti-CD3 antibody (clone OKT3, eBiosciences) or  $2.5 \mu\text{g ml}^{-1}$  PHA (Sigma-Aldrich) for 3 days or in the presence of candidin ( $5 \mu\text{g ml}^{-1}$ , Bio-Rad), tetanus toxoid (1 to 8,000 dilution, Statens Serum Institute) or PPD (tuberculin) ( $50 \mu\text{g ml}^{-1}$ , Statens Serum Institute) for 6 days. Then  $0.074 \text{ MBq ml}^{-1}$  of [ $^3\text{H}$ ]thymidine, [ $^3\text{H}$ ]cytidine, [ $^3\text{H}$ ]uridine or [ $^3\text{H}$ ]leucine or  $0.185 \text{ MBq ml}^{-1} \text{ U-}^{14}\text{C}$  aspartate were added during the last 18 h of stimulation. For [ $^3\text{H}$ ]cytidine, this corresponds to  $0.133 \mu\text{M}$ , which does not restore normal



proliferation of CTPS1-deficient cells that required 50  $\mu$ M. Cells were collected with a Filter Mate harvester (PerkinElmer) on filters for labelled cell assays (Uni-filter plates, PerkinElmer) that retain nucleic acids and filters were then washed. Radioactivity (c.p.m.) on filters (corresponding to radiolabelled compounds incorporated in nucleic acids) was measured by liquid scintillation counting with Top-Count NXT beta counter (PerkinElmer).

**Apoptosis assay.** Apoptosis was determined by evaluating phosphatidylserine (PS) exposure in the outer leaflet of the cytoplasmic membrane with PE-conjugated annexin-V in combination with 7-AAD Viaprobe (Apoptosis Detection Kit I, BD) 12 h after stimulation with coated-OKT3 (0.1, 1 or 10  $\mu$ g ml<sup>-1</sup>). Apoptosis was based on the percentage of annexin V<sup>+</sup>/7AAD<sup>-</sup> cells to exclude necrotic cells. Cells were analysed by flow cytometry.

**Cytokine production and degranulation.** For intracellular staining of cytokines, cells were stimulated overnight with PMA and ionomycin or anti CD3/CD28 beads in the presence of brefeldin A (GolgiPlug, BD). Cells were then fixed and permeabilized using the BD cytofix/cytoperm plus kit (BD Pharmingen) according to the manufacturer's instructions. Cells were labelled with PE-anti-IL-2 (rat IgG2a, MHQ1-17H12), PE/Cy7-anti-TNF- $\alpha$  (mouse IgG1; MAb11), APC-anti-IFN- $\gamma$  (mouse IgG1, 4S.B3) and isotype-matched monoclonal antibodies purchased from BioLegend and analysed by flow cytometry. Degranulation was determined by analysis of the expression of CD107/LAMP, a marker of the exocytosis of lytic granules. Blasts were stimulated for 3 h in the presence of 0.3, 3 or 30  $\mu$ g ml<sup>-1</sup> coated-OKT3 and simultaneously labelled with eFluor660-anti-CD107a (mouse IgG1; eBioH4A3), eFluor660-anti-CD107b (mouse IgG1; eBioABL-93) or isotype matched monoclonal antibodies purchased from eBiosciences. Thereafter, cells were collected, washed and stained with FITC-anti-CD3 and PE-anti-CD8 monoclonal antibodies and analysed by flow cytometry.

**Flow cytometry.** Cell staining and the flow-cytometry-based phenotypic analyses of PBMCs and cells were performed according to standard flow cytometry methods. The following monoclonal antibodies were conjugated to fluorescein isothiocyanate (FITC), phycoerythrin (PE), phycoerythrin-Cyanine7 (PE/Cy7), phycoerythrin-Cyanine5, phycoerythrin-Cyanine5.5, allophycocyanin (APC), allophycocyanin-Vio7, View blue or View green: anti-CD25 (M-A251), anti-CD27 (M-T271), anti-CD31 (WM59), anti-CD45RA (HI100), anti-CD45RO (UCHL1), anti-CD197/CCR7 (3D12), anti-TCR $\alpha\beta$  (T10B9), anti-TCR $\gamma\delta$  (B1), anti-CD95 (DX2), anti-CD19 (HIB19), anti-CD21 (B-ly4), anti-IgM (G20-127), anti-IgD (IA6-2), anti-CD56 (B159) and anti-CD16 (3G8), all purchased from BD Biosciences and anti-CD3 (BW264/56), anti-CD4 (VIT4), anti-CD8 (BW135/80) and anti-CD69 (FN50) from Miltenyi Biotec. iNKT cells were detected by staining with anti-V $\alpha$ 24 (C15) and anti-V $\beta$ 11 (C21) (Beckman Coulter) and MAIT cells by staining with anti-V $\alpha$ 7.2 (3C10) and anti-CD161 (HP-3G10) (Biolegend). All data were collected on a FACS-Canto II cytometer (BD Biosciences) and analysed using FlowJo Version 9.3.2 software (Tree Star).

**Immunoblotting and analysis of CTPS1 protein expression.** Cells ( $5 \times 10^6$  cells per ml) were stimulated by anti-CD3 antibody (1  $\mu$ g ml<sup>-1</sup>) crosslinking with a rabbit anti-mouse IgG (2  $\mu$ g ml<sup>-1</sup>) or anti CD3/CD28 beads for the indicated time periods. Cells were then lysed in 1% NP40, 50 mM Tris pH 8, 150 mM NaCl, 20 mM EDTA, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM NaF and complete protease inhibitor cocktail (Roche), as previously described<sup>30</sup>. Protein concentrations were quantitated by BCA assay (BIO-RAD). Then 80  $\mu$ g of proteins were separated by SDS-PAGE and transferred on PVDF membranes (Millipore). Membranes were blocked with milk or BSA before incubation with antibodies. The following monoclonal antibodies and rabbit polyclonal antibodies were used for immunoblotting: anti-PLC- $\gamma$ 1 (#2822S), anti-phosphorylated PLC- $\gamma$ 1 (#2821S), anti-phosphorylated ERK1/2 (clone E10, #9106S), anti-ERK1/2 (#4695S) anti-phosphorylated I $\kappa$ B $\alpha$  (clone 5A5, #9246S), anti-phosphorylated PKC $\theta$  (#9377S), NF $\kappa$ B (clone D14E12) anti-phosphorylated AKT (Serine 473, clone 587F11) and anti-phosphorylated tyrosine (4G10) purchased from Cell Signaling Technology and rabbit polyclonal antibodies anti-actin (#A2066) and anti-CTPS1 raised against the residues 341 to 355 (#SAB111071) or 416-430 (#SAB111072) and anti-CTPS2 (#HPA017437) purchased from Sigma-Aldrich. Anti-CTPS1 rabbit polyclonal antibodies (K-21) from Santa Cruz were also tested. Membranes were then washed and incubated with anti-mouse or anti-rabbit HRP-conjugated secondary antibodies from Cell Signaling and GE Healthcare, respectively. Pierce ECL western blotting substrate was used for detection. For inhibition assays of the signalling pathways after TCR-CD3 activation, cells have been stimulated with anti-CD3/CD28 beads for 48 h in the presence of 100 nM of the MAPK/ERK inhibitor PD0325901, 10  $\mu$ M of the Src family protein tyrosine kinase inhibitor PP1, 10  $\mu$ M of the Src family protein tyrosine kinase inhibitor PP2, 10  $\mu$ M of the selective Ca<sup>2+</sup> chelator 1,2-bis-(2-aminophenoxy) ethane-N,N,N',N'-tetraacetic acid tetra(acetoxymethyl) ester (BAPTA/AM), 10  $\mu$ M of the I $\kappa$ B $\alpha$  phosphorylation inhibitor Bay 11-7085 or 10  $\mu$ M of PI3Kdelta inhibitor IC87114. All were from Sigma-Aldrich, except IC87114 from Calbiochem. The concentrations used were typical and previously reported. After 48 h incubation with the different inhibitors, cell viability was verified and was more than 90% in each condition. The

activity and the selectivity of the inhibitors was verified in parallel by immunoblotting for phospho-tyrosine (for PP1 and PP2), I $\kappa$ B $\alpha$  phosphorylation (for Bay 11-7085), ERK phosphorylation (for PD0325901) and AKT phosphorylation (for IC87114) (data not shown).

**Calcium flux analysis.** Ca<sup>2+</sup> responses were assessed by flow cytometry, as previously described<sup>31</sup>. Briefly, cells were loaded with 5  $\mu$ M Indo-1 a.m. (Molecular Probes), washed, incubated with anti-CD4-APC and anti-CD8-PE monoclonal antibodies, stimulated by anti-CD3 antibody (0.125  $\mu$ g ml<sup>-1</sup>) crosslinking with F(ab')<sub>2</sub> rabbit anti-mouse IgG (10  $\mu$ g ml<sup>-1</sup>) and then incubated with ionomycin (1  $\mu$ M). Cells were analysed with a FACS Aria flow cytometer (BD Biosciences). Ca<sup>2+</sup> flux data were obtained using kinetic analyses of FlowJo software package (Tree Star). Intracellular Ca<sup>2+</sup> levels correspond to the normalized ratio of Ca<sup>2+</sup>-bound to Ca<sup>2+</sup>-free Indo-1 fluorescence and are plotted as a function of time.

**Plasmid constructs, cell transfections and infections.** A full-length cDNA encoding wild-type CTPS1 and a full length cDNA encoding the mutant CTPS1 $\Delta$ 18 were obtained by RT-PCR from control blasts and blasts from patient 1.2 respectively using the forward 5'-CGGGATCCACCATGAAGTATATTCTGGTT-3' and reverse 5'-CCGCTCGAGTCAGTCATGATTATTATGA-3' (for wild type) and 5'-CCGCTCGAGTTAAAGAAAGTCTCCAAGC-3' (for CTPS1 $\Delta$ 18) specific primers. The cDNAs were verified by sequencing and inserted into a bicistronic lentiviral expression vector encoding the green fluorescent protein (GFP) as a reporter (pLenti7.3/V5-TOPO, Invitrogen). Viral particles for infection were obtained by co-expression of the lentiviral vector containing CTPS1 with third-generation lentiviral plasmids containing Gag-Pol, Rev and the G protein of the vesicular stomatitis virus (VSV) into HEK 293T using calcium phosphate. Viral supernatants were collected every 12 h on 2 consecutive days, starting 48 h after transfection, and viral particles were concentrated by ultracentrifugation at 49,000g for 1.5 h at 12 °C. Cells were infected with viral particles at a minimal titre of 10<sup>7</sup> transducing units per ml and 48 h after infection, cells were deprived of IL-2 during 72 h for proliferation assays. To assess the selective advantage of GFP expression during long-term expansion, blasts were re-stimulated with anti-CD3/CD28 beads (Invitrogen) every 48 h during 8 days. For CTPS1 gene knockdown, blasts or Jurkat cells were infected at day 3 of PHA stimulation with the pLKO.1 lentiviral vector containing a CTPS1-specific shRNA (OpenBiosystems, n°TRCN0000045349 and n°TRCN0000045350) or a scrambled shRNA in which the puromycin resistance gene was replaced by the GFP gene. Proliferation was analysed in GFP<sup>+</sup> and GFP<sup>-</sup> blasts after 4 days of stimulation with anti-CD3/CD28 beads as previously described. For survey of loss of GFP expression in long-term expansions, blasts were repeatedly stimulated with anti-CD3/CD28 beads every 48 h during 8 days. Jurkat cells were maintained in culture after infection during 26 days. The proportions percentages of GFP<sup>+</sup> cells were determined by flow cytometry.

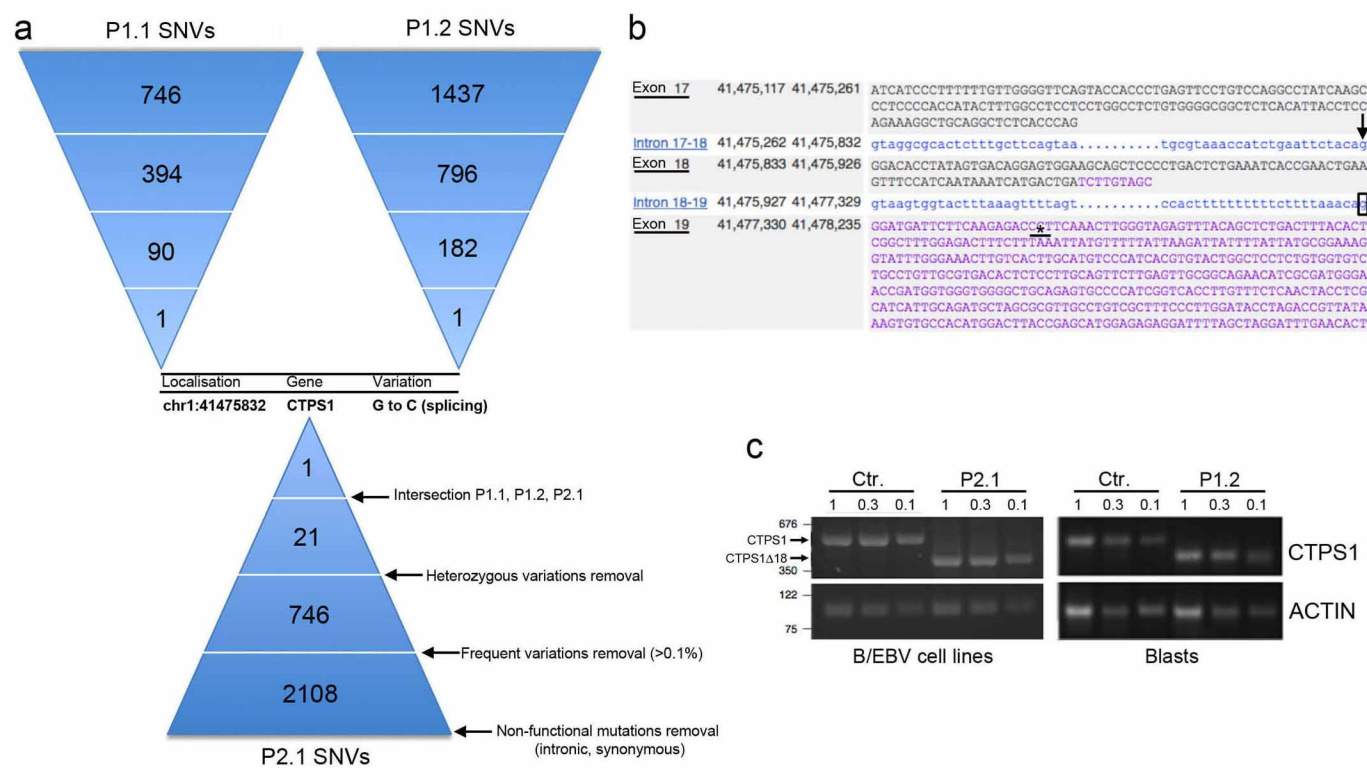
**Quantification of intracellular nucleotides.** Intracellular pools of nucleotides were quantified based on previously described methods<sup>32,33</sup>. Briefly, five million cells were washed in 0.1 M phosphate buffer (pH 7.4) and lysed in 60  $\mu$ l HClO<sub>4</sub> 1 M, containing 2  $\mu$ M 8-bromo-AMP (8-BrAMP) as an internal standard. After 12,000g centrifugation for 5 min at 4 °C, supernatants were transferred to a 384-well plate and kept at 4 °C in an auto sampler before injection. Aliquots of 5  $\mu$ l were injected onto a separation column (ACQUITY UPLC BEH300 C18, 1.7  $\mu$ m, 2.1  $\times$  100 mm reversed-phase column, Waters) with a flow rate of 0.5 ml min<sup>-1</sup> and analysed with a tandem mass spectrometry system consisting of an Acquity Ultra Performance Liquid Chromatography (UPLC Waters) interfaced with a xevo-TQ-S tandem quadrupole mass spectrometer (Waters). Mobile phase A was 0.1% formic acid in water and mobile phase B, 0.1% formic acid in acetonitrile. A programmed mobile phase-gradient was used during a 7-min run: 0 min, 1% B; 5 min, 10% B; 5.1 min, 100% B; 6 min, 100% B; 6.1 min, 1% B; 7 min, 1% B. The content of the 4 nucleotides ATP, GTP, UTP and CTP was quantified in the electrospray negative ion mode with multiple reaction monitoring (MRM). Transitions of  $m/z$  505.9 > 408 and 505.9 > 272.9 were used for quantification and confirmation of ATP, respectively, and those of 521.9 > 158.9 and 521.9 > 177 for GTP, 482.8 > 158.9 and 482.8 > 79 for UTP, and 481.8 > 158.9 and 481.8 > 384 for CTP. Concentrations were determined by using calibration curves of the 4 nucleotides. The linearity, exactitude and variability were determined for the technical validation of this assay. The linearity gave a correlation coefficient of the linear regression curves greater than 0.99 for the 4 nucleotides. The minimum and maximum recovery of spiked samples with the 4 nucleotides at a concentration of 90 mg l<sup>-1</sup> and 250 mg l<sup>-1</sup> ranged from 72% to 123%. The maximum intra- and inter-assay variability was 22% and 23%, respectively.

**Statistical analysis.** *P* values were calculated with a Student's *t*-test using PRISM software (GraphPad Software), with a two-tailed distribution. The variance was similar between the groups that have been statistically compared.

29. Gülow, K. *et al.* HIV-1 trans-activator of transcription substitutes for oxidative signaling in activation-induced T cell death. *J. Immunol.* **174**, 5249–5260 (2005).
30. Latour, S. *et al.* Regulation of SLAM-mediated signal transduction by SAP, the X-linked lymphoproliferative gene product. *Nature Immunol.* **2**, 681–690 (2001).

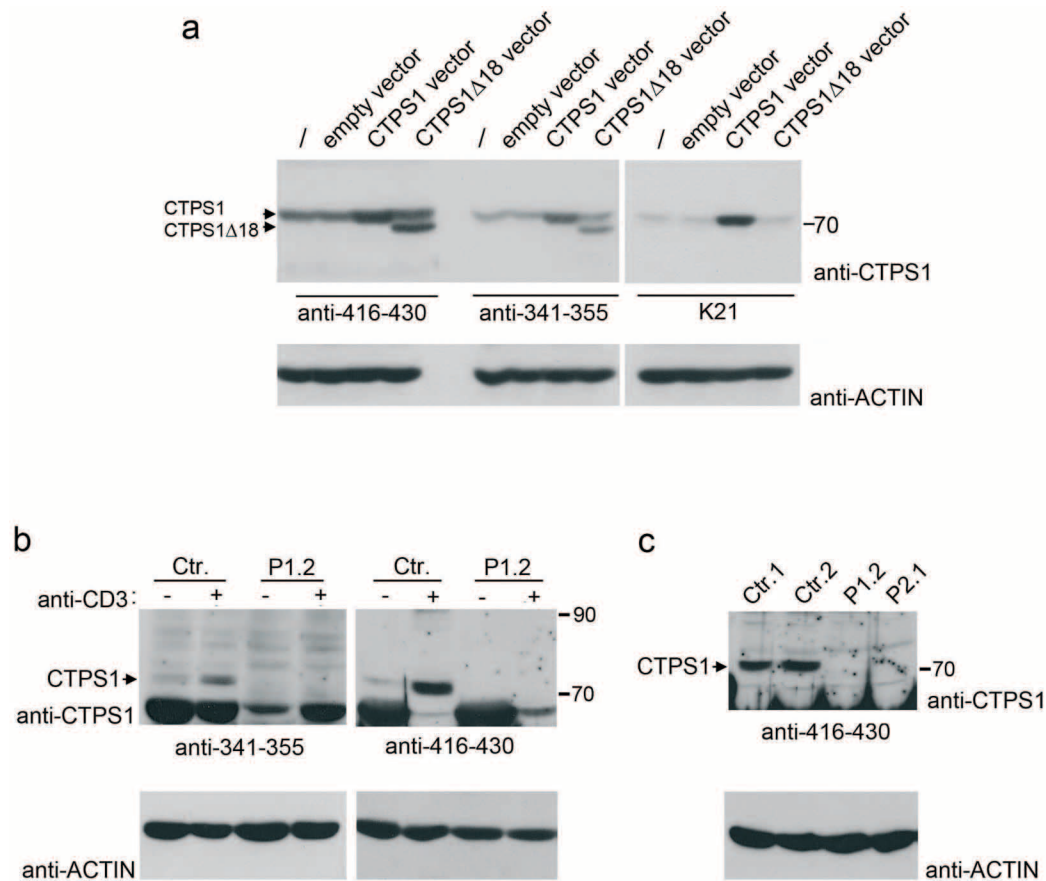


31. Picard, C. *et al.* Hypomorphic mutation of *ZAP70* in human results in a late onset immunodeficiency and no autoimmunity. *Eur. J. Immunol.* **39**, 1966–1976 (2009).
32. Luo, B., Groenke, K., Takors, R., Wandrey, C. & Oldiges, M. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography-mass spectrometry. *J. Chromatogr. A* **1147**, 153–164 (2007).
33. Scavennec, J., Maraninchi, D., Gastaut, J. A., Carcassone, Y. & Cailla, H. L. Purine and pyrimidine ribonucleoside monophosphate patterns of peripheral blood and bone marrow cells in human acute leukemias. *Cancer Res.* **42**, 1326–1330 (1982).



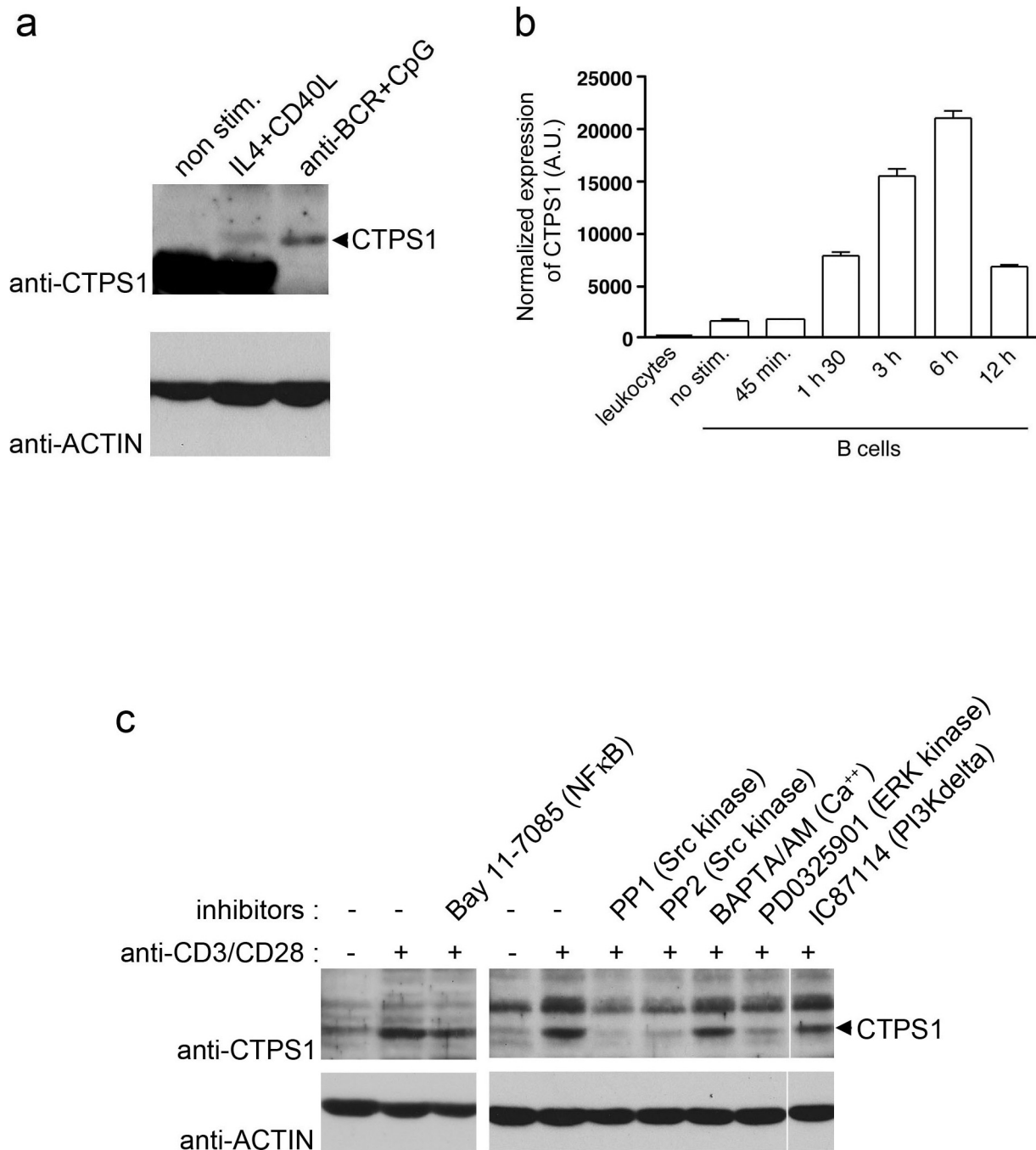
**Extended Data Figure 1 | Identification of a genetic *CTPS1* defect in patients P1.1, P1.2 and P2.1.** **a**, Analysis of the single nucleotide variations (SNVs) detected by whole-exome sequencing in the genome of P1.1, P1.2 and P2.1. The numbers of SNVs are indicated in the triangles. SNVs were filtered by removal of non-functional intronic and synonymous mutations, heterozygous variations and those present in dbSNPs, 1000 genomes databases. The intersection of the filtered SNVs in the three patients resulted in the identification of a single common splicing site variation in the *CTPS1* gene. **b**, Exon-intron structure and sequences of exons 17, 18 and 19 of *CTPS1*. The position of the variation is indicated by an arrow. The boxed nucleotide corresponds to the alternative splice site which produces a shorter transcript

lacking exon 18 detected in patient cells. The alternative stop codon is indicated by an asterisk. **c**, Expression of a *CTPS1* transcript lacking exon 18 (*CTPS1*Δ18) in *CTPS1*-deficient patients. The relative expression of full length *CTPS1*, *CTPS1*Δ18 and actin transcripts was examined by qRT-PCR in EBV-B cell lines (patient P2.1) and T-cell blasts (patient P1.2) from *CTPS1*-deficient patients. qRT-PCRs of actin are shown as normalization controls of the cDNA samples. Three fold-serial dilutions of cDNAs (indicated as 1, 0.3 and 0.1) were used for amplification of each transcript. Base pair markers are shown on the left. PCR products were verified by sequencing showing the expression of an abnormal *CTPS1* transcript lacking exon 18 in the cells of the patients.



**Extended Data Figure 2 | Loss of CTPS1 expression and undetectable expression of the mutant CTPS1Δ18 protein in cells from CTPS1-deficient patients.** **a**, Transient expression of CTPS1 and the mutant CTPS1Δ18 in 293-T cells transfected with vectors containing wild-type CTPS1 or the mutant CTPS1Δ18. Cell lysates were tested by immunoblotting for CTPS1 with different antibodies raised against CTPS1 and for actin as a control for loading. The CTPS1Δ18 mutant protein is recognized by the rabbit polyclonal antibodies raised against the 341 to 355 (anti-341-355) or the 416 to 430

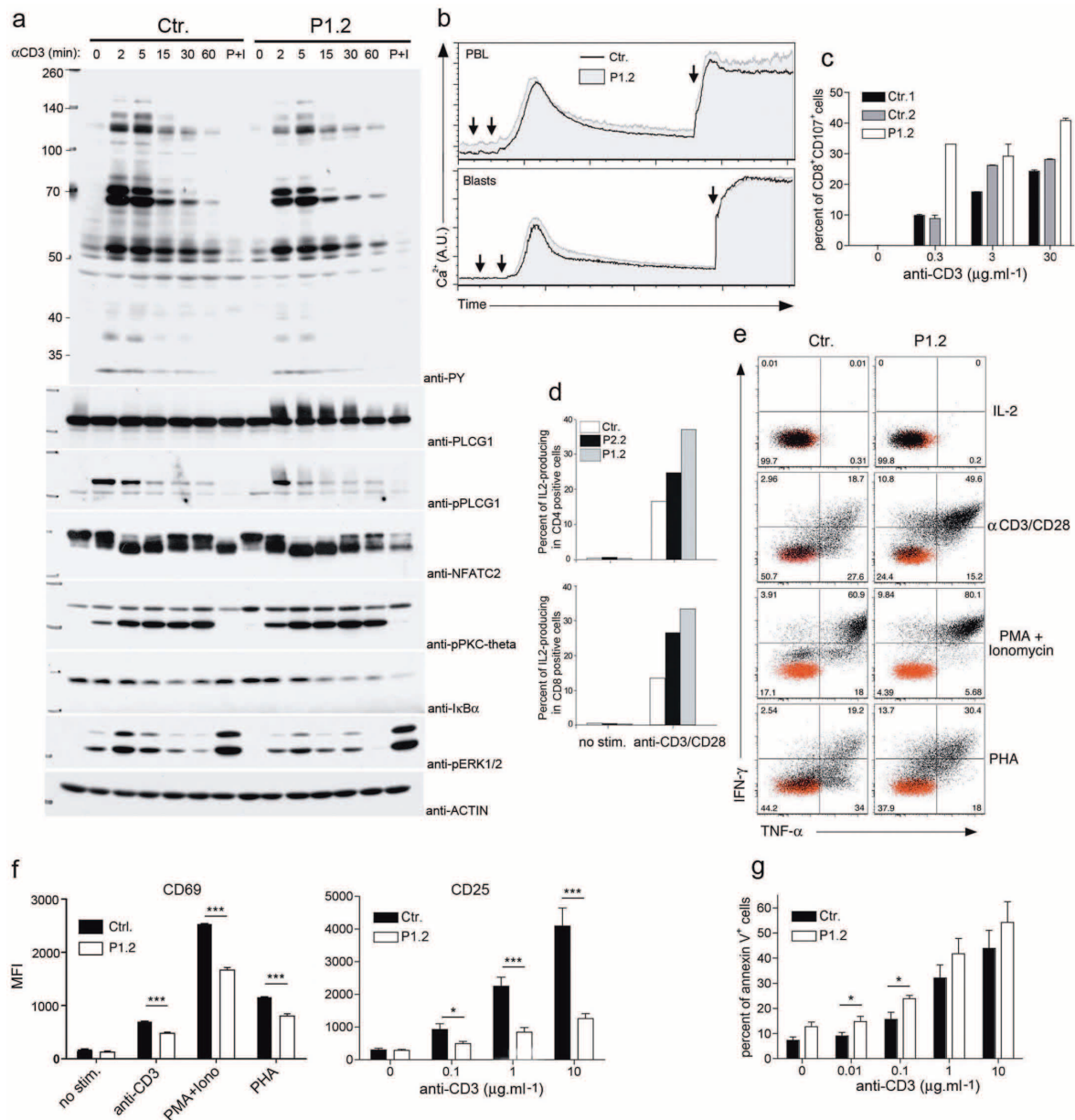
(anti-416-430) residues of CTPS1 but not by the rabbit polyclonal antibody K21. **b**, T-cell blasts from a healthy control (Ctr.) and the CTPS1-deficient patient P1.2 (P1.2) stimulated for 48 h with anti-CD3 were analysed for CTPS1 expression with the rabbit polyclonal antibodies anti-416-430 and anti-341-355. Actin expression as control for loading. **c**, EBV B-cell lines from healthy controls (Ctr. 1 and Ctr.2) and CTPS1-mutated patients (P1.2 and P2.1) were analysed for CTPS1 expression with the rabbit polyclonal antibody anti-416-430. Actin expression served as control for loading.



**Extended Data Figure 3 | Induction of CTPS1 expression in activated B cells.** **a**, Immunoblots for CTPS1 expression in sorted CD19<sup>+</sup> B cells (from PBMCs of an healthy donor) stimulated with the indicated stimuli. Actin was used as a loading control. **b**, Kinetics of *CTPS1* mRNA expression monitored by qRT-PCR in sorted B cells that have been stimulated with anti-BCR+CpG. Expression is in arbitrary units (a.u.) normalized to the

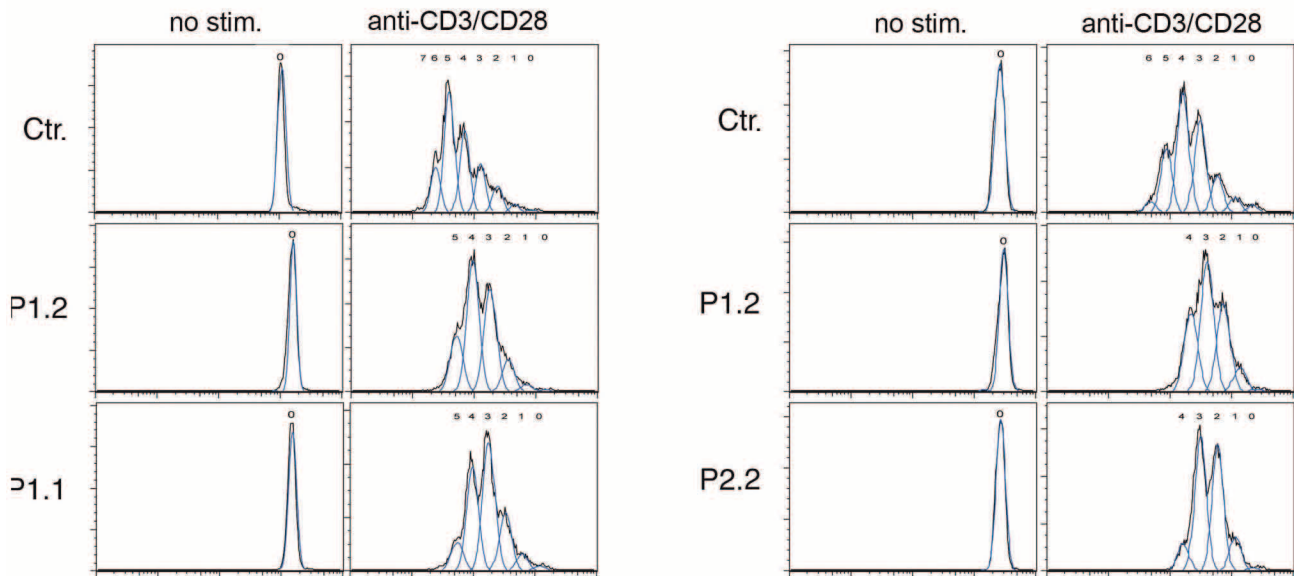
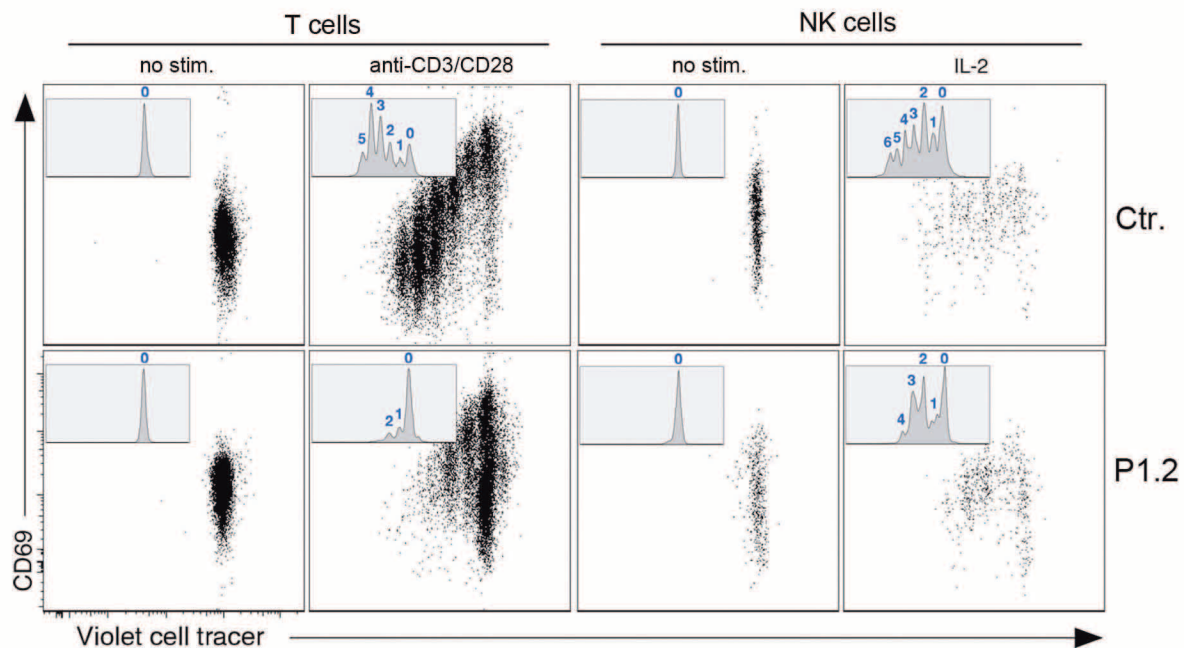
expression of the *GADPH* gene and leukocytes were used as calibrator. **c**, Immunoblots for CTPS1 expression in T-cell blasts (from an healthy donor) stimulated with anti-CD3/CD28 beads in the presence of selective inhibitors of NFκB, Src kinases, Ca<sup>2+</sup>, ERK kinase and PI3Kdelta. Actin was used as a loading control. The activity of the inhibitors was controlled in parallel (see Methods and data not shown).





**Extended Data Figure 4 | Analysis of proximal and late TCR activation responses in CTPS1-deficient cells.** **a**, Immunoblots showing the phosphorylation of proximal signalling molecules in T-cell blasts from a control donor (Ctr.) and a CTPS1-deficient patient P1.2 (P1.2) stimulated with anti-CD3 antibodies for 0, 2, 5, 15, 30 and 60 min or PMA plus ionomycin (P + I). Cell lysates were immunoblotted with antibodies against tyrosine-phosphorylated residues (PY), phosphoPLCG1 (pPLCG1), PLCG1, NFATc2, phosphoPKCtheta (pPKCtheta), IκBα, phosphoERK1/2 (pERK1/2) and actin as a loading control. Molecular weights are on the left. Data correspond to one representative experiment of 2 or 3 independent experiments. **b**, Flow cytometry analyses of  $\text{Ca}^{2+}$ -flux in T cells from PBMCs or T-cell blasts of a control donor (Ctr.) and a CTPS1-deficient patient P1.2 (P1.2) loaded with the  $\text{Ca}^{2+}$ -sensitive fluorescent dye Indo-1. Cells were then stimulated with anti-CD3 antibodies (first arrow) crosslinked with rabbit anti-mouse antibodies (second arrow) and then incubated with ionomycin (third arrow) to induce a receptor-independent  $\text{Ca}^{2+}$  response. Intracellular  $\text{Ca}^{2+}$  levels are expressed in arbitrary units (a.u.). Data with the T-cell blasts correspond to one of three representative experiments. **c**, Analysis of the degranulation capacity of  $\text{CD8}^{+}$  T-cell blasts from two control donors (Ctr.1 and Ctr.2) and a CTPS1-deficient patient P1.2 (P1.2) stimulated with the indicated concentrations of anti-CD3 antibodies for 4 h. Cells were stained with antibodies against CD107a/b (LAMP1/2), a surface-exposed marker of the secretion of lytic granules, and then analysed by FACS. Means with s.d. of percentages of  $\text{CD8}^{+}$

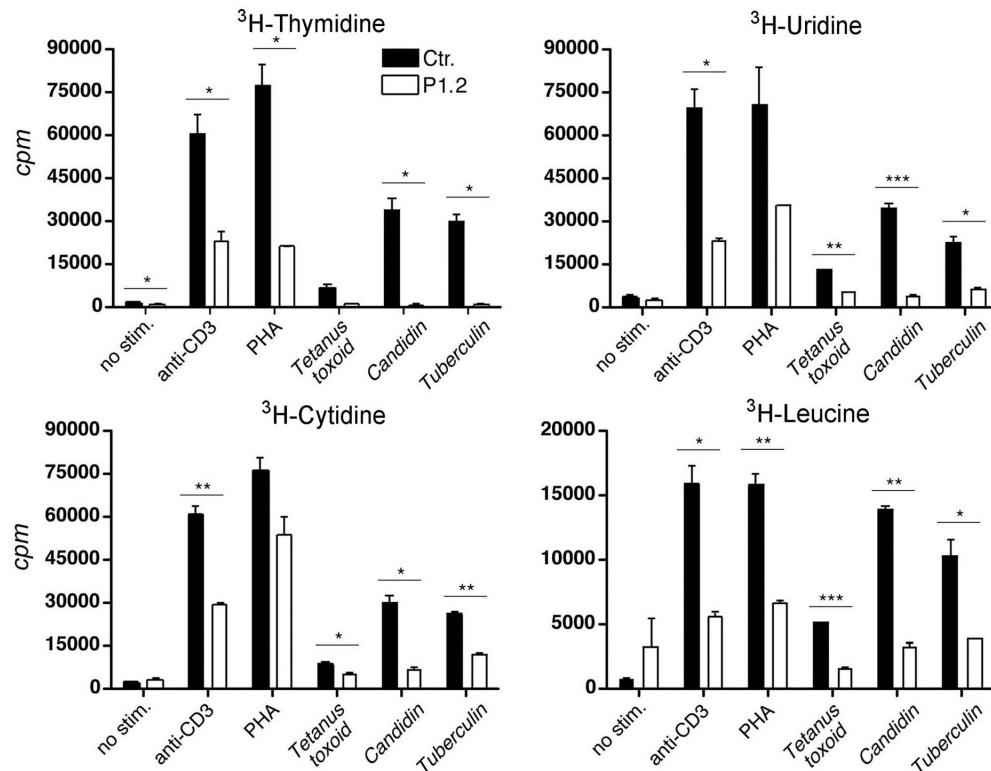
$\text{CD107}^{+}$  cells are presented. **d**, Flow cytometry analysis of intracellular IL-2 production in  $\text{CD4}^{+}$  and  $\text{CD8}^{+}$  T cells from PBMCs of a control donor (Ctr.) and two CTPS1-deficient patients P1.2 and P2.2 (P1.2 and P2.2) stimulated for 36 h with anti-CD28 and anti-CD3 antibodies. The percentages of  $\text{CD4}^{+}\text{IL-2}^{+}$  and  $\text{CD8}^{+}\text{IL-2}^{+}$  are shown. **e**, Flow cytometry analysis of intracellular IFN- $\gamma$  and TNF- $\alpha$  production on gated  $\text{CD3}^{+}$  T-cell blasts of a control donor (Ctr.) and a CTPS1-deficient patient P1.2 (P1.2) stimulated for 12 h with IL-2, anti-CD3 and anti-CD28 coated beads (anti-CD3/CD28), PMA plus ionomycin or PHA. Data are representative of one of 3 independent experiments. Dot-plots in red correspond to the isotype control. **f**, Induction of CD25 and CD69 in  $\text{CD3}^{+}$  T-cell blasts from a control donor (Ctr.) and a CTPS1-deficient patient (P1.2) was assessed after 24 h of anti-CD3 stimulation for CD69 and 96 h for CD25. Expression was assessed by flow cytometry and the median fluorescence intensity (MFI) is presented. Data are means with s.d. of four and eight independent experiments for CD69 and CD25, respectively. Unpaired Student's *t*-test. \*\*\* $P < 0.001$ , \* $P < 0.05$ . **g**, Analysis of activation-induced cell death (AICD) in  $\text{CD3}^{+}$  T-cell blasts from a control donor (Ctr.) and a CTPS1-deficient patient P1.2 (P1.2) after stimulation with the indicated concentration of anti-CD3 antibodies for 12 h. Apoptotic cells were detected by annexin V and 7-AAD staining and the percentages of annexin V positive/7-AAD negative cells within the gated CD3 population are shown. Data are means with s.d. of four (P1.2,  $n = 4$ ) and eight (Ctr.,  $n = 8$ ) independent experiments. Unpaired Student's *t*-tests and \* $P < 0.05$ .

**a****b**

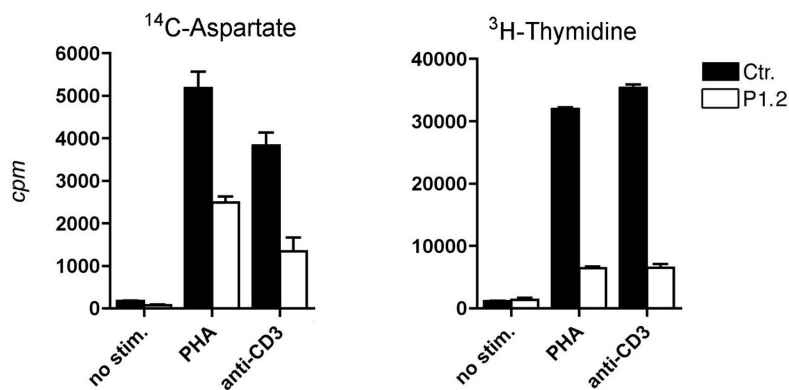
**Extended Data Figure 5 | Decreased proliferation of TCR-stimulated T cells from patients P1.1, P1.2 and P2.2, and IL-2-expanded natural killer cells from patient P1.2.** **a**, Proliferation of  $CD3^+$  T cells from PBMCs of control donors (Ctr.) and CTPS1-deficient patients (P1.1, P1.2; left panels) or (P1.2, P2.2; right panels). Right panels and left panels correspond to 2 independent experiments. Cells were stimulated with immobilized anti-CD3 and soluble anti-CD28 antibodies during the course of 6 days. The proliferation was determined by dilution of CFSE staining analysed by flow cytometry.

Histograms correspond to CFSE staining dilutions for which the number of cell divisions was indicated at the top of each peak. **b**, Proliferation of  $CD3^+$  T and  $CD16^+CD56^+$  natural killer cells from PBMCs of a control donor (Ctr.) and a CTPS1-deficient patient (P1.2). Cells were stimulated with anti-CD3/CD28 coated beads for 3 days or IL-2 for 7 days. Representative dot plots showing cell divisions by dilution of the violet dye and expression of the activation marker CD69. Inserts with histograms showing the violet dye dilution with the number cell divisions indicated at the top of each peak.

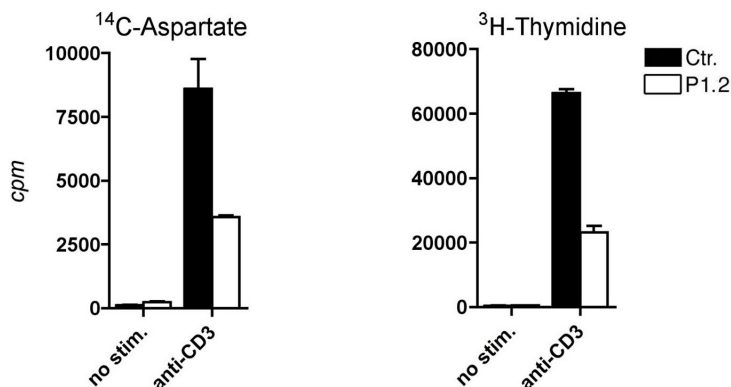
a



b



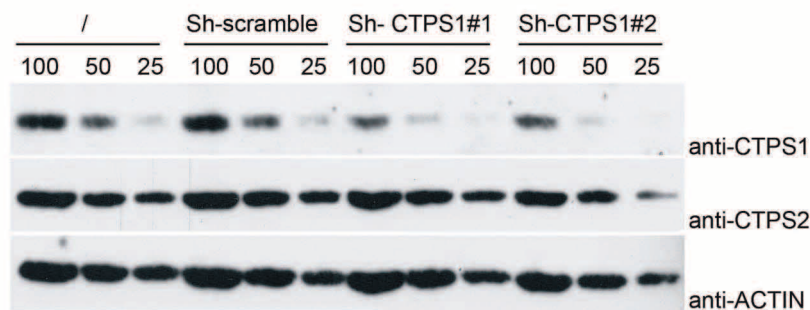
c



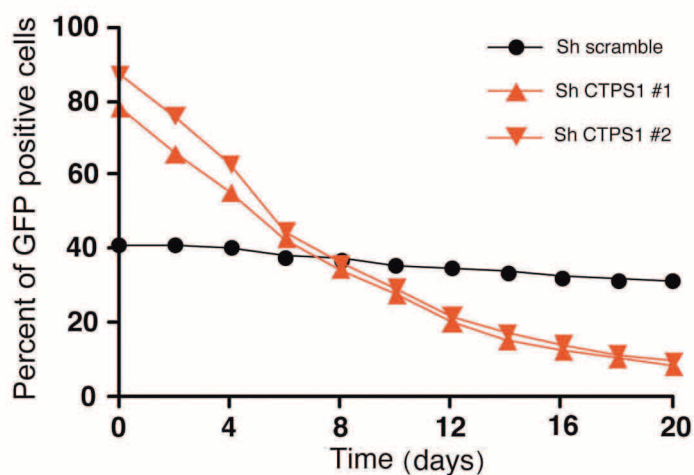
**Extended Data Figure 6 | Decreased incorporation of thymidine, uridine, cytidine, leucine and aspartate in CTPS1-deficient cells.** a, Incorporation of [ $^3\text{H}$ ]thymidine, [ $^3\text{H}$ ]uridine, [ $^3\text{H}$ ]cytidine and [ $^3\text{H}$ ]leucine as tracers of DNA, RNA and protein synthesis in PBMCs from a control healthy donor (Ctr.) and a CTPS1-deficient patient (P1.2) stimulated or not (no stim.) for 3 days with anti-CD3 or PHA and for 6 days with tetanus toxoid, candidin or tuberculin. The concentration of [ $^3\text{H}$ ]cytidine used in these experiments is under the value

allowing the restoration of normal proliferation in CTPS1-deficient cells (also see Methods). Data are means with s.d. of two independent experiments with triplicates. Unpaired Student's *t*-tests and \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ . b, c, Incorporation of [ $^{14}\text{C}$ ]aspartate and [ $^3\text{H}$ ]thymidine in PBMCs (b) or T-cell blasts (c) from a control healthy donor (Ctr.) and a CTPS1-deficient patient (P1.2) stimulated or not (no stim.) for 3 days with anti-CD3 or PHA. Data means with s.d. of three independent samples.

a



b

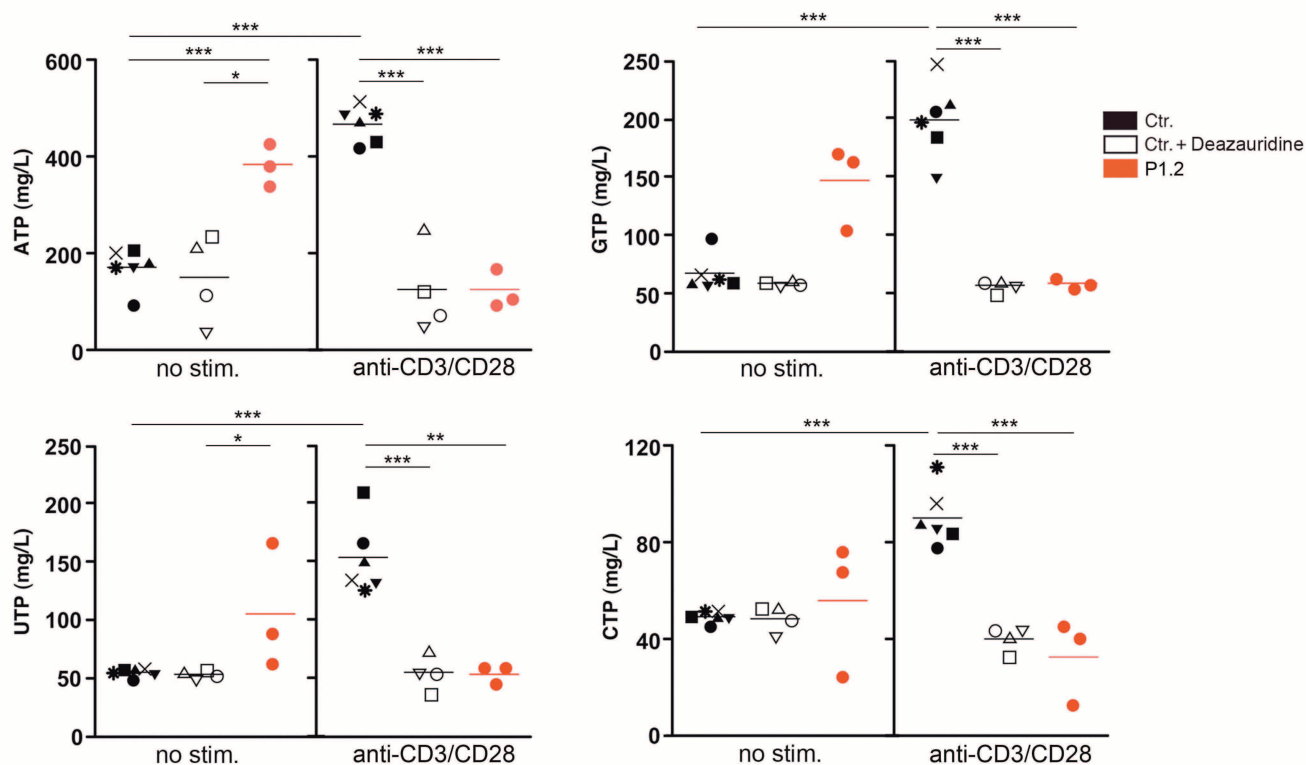


**Extended Data Figure 7 | Decreased proliferation of Jurkat cells following shRNA-mediated CTPS1 downregulation.** **a**, Expression of CTPS1 in Jurkat cells transduced with lentiviral vectors containing two distinct CTPS1 shRNAs (Sh CTPS1 #1 or Sh CTPS1 #2) or a scrambled shRNA (Sh scramble). Cell lysates were analysed by immunoblotting for CTPS1, CTPS2 and actin protein

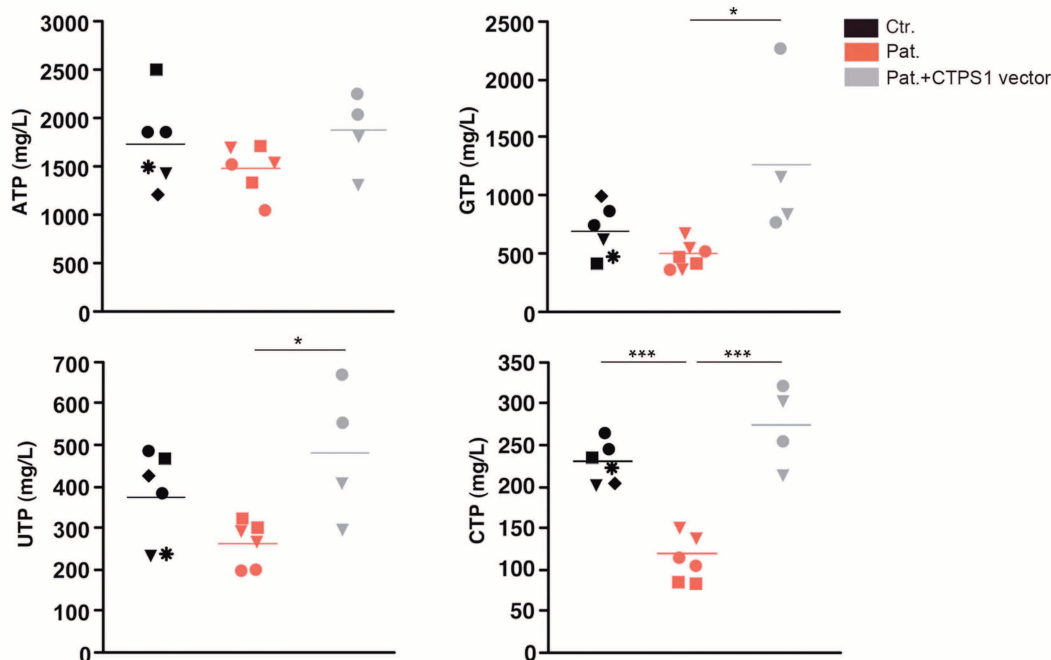
expression. **b**, Proliferation of Jurkat cells, in which CTPS1 expression was silenced, was monitored as a function of the loss of GFP expression and compared with cells transduced with the scrambled shRNA. The percentages of GFP-positive cells were determined by flow cytometry at the indicated time points with 'time 0' corresponding to 48 h post-transduction.



a



b



#### Extended Data Figure 8 | Measurements of nucleotide pools in CTPS1-deficient T cells and in B/EBV cell lines from patients.

**a**, Concentration of ATP, GTP, UTP and CTP in cell extracts of T-cell blasts stimulated with anti-CD3/CD28 coated beads or not (no stim.) from control healthy donors (Ctr.) or from patient P1.2. Control cells were treated or not with deazauridine for 24 h before and during stimulation or not. Representative data from 3 independent experiments. **b**, Same as in **a** with EBV B-cell lines

from control healthy donors (Ctr.) and patients P1.1 (squares), P1.2 (circles) and P2.1 (triangles). For controls, each symbol corresponds to a different control cell line (from a different healthy donor). Representative data from two independent experiments with blinding during the measurements. Bars correspond to averages. Unpaired Student's *t*-tests and \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001. CTP data are also shown in Fig. 3g, h.

Extended Data Table 1 | Immunological data of CTPS1-deficient patients

Patient	P1.1	P1.2	P2.1	P2.2	P3.1*	P3.2*	P4	P5**
Age (months)	38	113	110	48	60	20	24	60
<b>Cell subsets (cells.mm<sup>-3</sup>)</b>								
PMN (1800-8000)	2300	1100		4500	1100		6400	<b>450</b>
Lymphocytes (1700-6900)	3500	3400	<b>400</b>	<b>1600</b>	<b>200</b>	2400	4000	<b>440</b>
CD3 <sup>+</sup> (900-4500)	4390	1680	<b>280</b>	990	<b>80</b>	1461	2160	<b>340</b>
CD4 <sup>+</sup> (500-2400)	1410	680	<b>160</b>	570	<b>43</b>	<b>1095</b>	1320	<b>314</b>
CD8 <sup>+</sup> (300-1600)	2580	910	<b>120</b>	480	<b>28</b>	<b>346</b>	880	<b>35</b>
CD19 <sup>+</sup> (200-2100)	1010	540	<b>60</b>	<b>170</b>	<b>12</b>	297	760	<b>13</b>
CD56 <sup>+</sup> (100-1000)		480	<b>30</b>	120	<b>20</b>	128		<b>16</b>
<b>T-cell proliferation</b>								
PHA	<b>low</b>	normal		normal				normal
<b>Serum immunoglobulins (g.L<sup>-1</sup>)</b>								
IgG	24.7 (6.1-10.7)	29.5 (8.3-14.3)	10.6 (8.3-14.3)	<b>5.53</b> (6.8-11.8)	12.8 (6.8-11.8)	24.5 (5.3-10.1)	17.9 (6.8-11.8)	4.8 (4.2-8)
IgG2	<b>&lt;0.1</b> (4)		<b>0.23</b> (4)	<b>0.2</b> (4)				
IgA	0.6 (0.5-1.1)	1.72 (1-1.9)	0.74 (1-1.9)	<b>0.42</b> (0.7-1.3)	8.73 (0.7-1.3)	1 (0.3-0.8)	1.5 (0.7-1.3)	4.4 (0.2-0.7)
IgM	0.84 (0.5-1.1)	<b>0.37</b> (0.7-1.3)	0.72 (0.7-1.3)	<b>0.42</b> (0.5-1.1)	0.76 (0.5-1.1)	1.4 (0.5-1.1)	1.2 (0.5-1.1)	0.8 (0.5-1.1)
<b>Specific antibodies (IU.ml<sup>-1</sup>)</b>								
Tetanus (> 0.1)	0.16		0.48	<b>0.05</b>	1.44		<b>0.03</b>	0.57
<i>H. influenzae</i> type B (>1)	7.0		<b>0.74</b>	<b>0.5</b>	<b>0.86</b>			<b>0.06</b>
<i>S. Pneumoniae</i> (> 0.35)	<b>low</b>		protective	<b>low</b>	<b>low</b>		<b>low</b>	<b>low</b>

PMN, polymorphonuclear neutrophils. PHA, phytohaemagglutinin. Ig, immunoglobulin.

In bold correspond to values below normal age-matched ranges that are indicated in brackets.

\*when acutely unwell with VZV.

\*\*at presentation with EBV-driven LPD

Different immunological parameters of patients were tested from blood (numbers of cells), PBMCs (proliferation in response to PHA evaluated by incorporation of [<sup>3</sup>H]thymidine) and serum (immunoglobulin subclasses and specific antibodies).

Extended Data Table 2 | Immunological features of PBMCs from patient P1.2 at 8 and 9 years

	Normal values (age-matched)	8 years	9 years
Lymphocytes (cells.mm <sup>-3</sup> )	(1900-3700)	1700	1800
<b>T cells</b>			
CD3 <sup>+</sup> (cells.mm <sup>-3</sup> )	(1200-2600)	1207	1188
CD4 <sup>+</sup> (cells.mm <sup>-3</sup> )	(650-1500)	527	396
CD8 <sup>+</sup> (cells.mm <sup>-3</sup> )	(370-1100)	612	720
CD4/CD8 ratio	(0.9-2.6)	0.86	0.48
TCRα/β (%)	(26-85)	74	75
TCRγ/δ (%)	(0.2-14)	1.3	1
CD31 <sup>+</sup> CD45RA <sup>+</sup> / CD4 <sup>+</sup> (recent naïve thymic emigrant) (%)	(43-55)	47	39
CD45RO <sup>+</sup> / CD4 <sup>+</sup> (memory) (%) (cells.mm <sup>-3</sup> )	(13/30) (85/450)	44 255	50 198
CCR7 <sup>+</sup> CD45RA <sup>+</sup> / CD8 <sup>+</sup> (naïve) (%)	(52/68)	24	8
CCR7 <sup>+</sup> CD45RA <sup>-</sup> / CD8 <sup>+</sup> (central memory) (%)	(3-4)	3	1
CCR7 <sup>-</sup> CD27 <sup>-</sup> CD45RA <sup>-</sup> / CD8 <sup>+</sup> (effector memory) (%) (cells.mm <sup>-3</sup> )	(11/20) (42/220)	64 391	81 583
CCR7 <sup>+</sup> CD27 <sup>-</sup> CD45RA <sup>+</sup> / CD8 <sup>+</sup> (exhausted effector memory - EMRA) (%)	(1-18)	3.8	3.9
CD127 <sup>low</sup> CD25 <sup>+</sup> / CD4 <sup>+</sup> (regulatory) (%)	(2-8)	3.92	4.9
Vα7 <sup>+</sup> CD161 <sup>+</sup> / CD3 <sup>+</sup> (MAIT) (%)	(1-8)	0	0
Vα24 <sup>+</sup> Vβ11 <sup>+</sup> CD161 <sup>+</sup> / CD3 <sup>+</sup> (NKT) (%)	(>0.02)	0	0
<b>T-cell proliferation (cpm.10<sup>-3</sup>)</b>			
PHA (6.25 mg ml <sup>-1</sup> )	(>50)	84	27
OKT3 (50 ng ml <sup>-1</sup> )	(>30)	N.D.	26
<i>Candidine</i>	(>10)	8,4	3.1
<i>Tetanus toxoid</i>	(>10)	6,3	1.2
<i>Tuberculin</i>	(>10)	4,2	1
<b>NK cells</b>			
CD16 <sup>+</sup> CD56 <sup>+</sup> (cells.mm <sup>-3</sup> )	(100-480)	68	162
CD16 <sup>+</sup> CD56 <sup>+</sup> (%)	(4-17)	4	9
<b>B cells</b>			
CD19 <sup>+</sup> (cells.mm <sup>-3</sup> )	(270-860)	391	378
CD19 <sup>+</sup> (%)	(13-27)	23	9
CD21 <sup>+</sup> CD27 <sup>+</sup> / CD19 <sup>+</sup> (memory) (%)	(11-24)	3	3.2
IgD <sup>+</sup> IgM <sup>+</sup> / CD19 <sup>+</sup> CD21 <sup>+</sup> CD27 <sup>+</sup> (marginal zone) (%)	(31-51)	53.3	50.6
IgD <sup>-</sup> IgM <sup>-</sup> / CD19 <sup>+</sup> CD21 <sup>+</sup> CD27 <sup>+</sup> (switched) (%)	(21-49)	35.4	23.2

Different T-cell subsets, B-cell subsets and natural killer cells from PBMCs were tested by flow cytometry. T cell proliferation from PBMCs in response to different stimuli including antigen- specific responses was analysed.

# A Ctf4 trimer couples the CMG helicase to DNA polymerase $\alpha$ in the eukaryotic replisome

Aline C. Simon<sup>1\*</sup>, Jin C. Zhou<sup>2\*</sup>, Rajika L. Perera<sup>1†</sup>, Frederick van Deursen<sup>3</sup>, Cecile Evrin<sup>4</sup>, Marina E. Ivanova<sup>1†</sup>, Mairi L. Kilkenny<sup>1</sup>, Ludovic Renault<sup>2</sup>, Svend Kjaer<sup>5</sup>, Dijana Matak-Vinković<sup>6</sup>, Karim Labib<sup>4</sup>, Alessandro Costa<sup>2</sup> & Luca Pellegrini<sup>1</sup>

**Efficient duplication of the genome requires the concerted action of helicase and DNA polymerases at replication forks<sup>1</sup> to avoid stalling of the replication machinery and consequent genomic instability<sup>2–4</sup>. In eukaryotes, the physical coupling between helicase and DNA polymerases remains poorly understood. Here we define the molecular mechanism by which the yeast Ctf4 protein links the Cdc45–MCM–GINS (CMG) DNA helicase to DNA polymerase  $\alpha$  (Pol  $\alpha$ ) within the replisome. We use X-ray crystallography and electron microscopy to show that Ctf4 self-associates in a constitutive disk-shaped trimer. Trimerization depends on a  $\beta$ -propeller domain in the carboxy-terminal half of the protein, which is fused to a helical extension that protrudes from one face of the trimeric disk. Critically, Pol  $\alpha$  and the CMG helicase share a common mechanism of interaction with Ctf4. We show that the amino-terminal tails of the catalytic subunit of Pol  $\alpha$  and the Sld5 subunit of GINS contain a conserved Ctf4-binding motif that docks onto the exposed helical extension of a Ctf4 protomer within the trimer. Accordingly, we demonstrate that one Ctf4 trimer can support binding of up to three partner proteins, including the simultaneous association with both Pol  $\alpha$  and GINS. Our findings indicate that Ctf4 can couple two molecules of Pol  $\alpha$  to one CMG helicase within the replisome, providing a new model for lagging-strand synthesis in eukaryotes that resembles the emerging model for the simpler replisome of *Escherichia coli*<sup>5–8</sup>. The ability of Ctf4 to act as a platform for multivalent interactions illustrates a mechanism for the concurrent recruitment of factors that act together at the fork.**

Recent evidence indicates that the leading- and lagging-strand polymerases are anchored to the helicase by replisome components that lack counterparts in bacteria<sup>9–13</sup>. The yeast Ctf4 protein is among the best characterized of these factors: its role is to link the CMG helicase with Pol  $\alpha$ , the polymerase subunit of the Pol  $\alpha$ –primase complex that initiates Okazaki fragments during lagging-strand synthesis<sup>9,10</sup>. Ctf4 is part of a conserved family of replication factors that includes human AND-1 and fission yeast Mcl1, and is required for efficient DNA synthesis, normal cell-cycle progression and genomic stability<sup>13–18</sup>. In addition to their role in DNA replication, Ctf4 and AND-1 perform an important yet poorly understood function in sister chromatid cohesion<sup>19–22</sup>.

Earlier work had shown that Ctf4 binds directly to the GINS subunit of the CMG helicase and to the catalytic subunit of Pol  $\alpha$ , via the C-terminal half of the protein that does not include an annotated WD40 domain in the N terminus of Ctf4 (ref. 10). We identified by bioinformatic analysis a second WD40 domain in the C-terminal half of yeast Ctf4, juxtaposed to a predicted helical region. Crystallographic analysis of residues 471–927 (C-end; 457 amino acids) of yeast Ctf4 (Ctf4(CTD)) confirmed the presence of a six-bladed  $\beta$ -propeller domain fused to a helical bundle of six  $\alpha$ -helices arranged in a stack of helical hairpins (Extended Data Table 1 and Extended Data Fig. 1). Notably, the structural analysis revealed a trimeric assembly of Ctf4 molecules, resulting from side-on packing of  $\beta$ -propeller domains (Fig. 1a). The homotypic

association of the  $\beta$ -propeller domains generates a discoidal shape with three-fold symmetry. The helical domains of each Ctf4 protomer extend upwards and away from the plane of the trimer, like legs of a three-legged stool. The trimeric assembly seems to be constitutive as it buries a total surface area of 8,100 Å<sup>2</sup>, an average of 2,700 Å<sup>2</sup> per interface. The existence of Ctf4 as a constitutive trimer mediated by self-association of its C-terminal domain was confirmed by single-particle electron microscopy (EM) analysis, which showed the presence of three-fold symmetry in particles of full-length Ctf4 and Ctf4(CTD) (Fig. 1b, c and Extended Data Fig. 2). The EM analysis of full-length Ctf4 further revealed that the N-terminal WD40 domains depart radially from the Ctf4(CTD) trimer, to which they are loosely connected (Fig. 1b). The presence of Ctf4 as a stable trimer in solution was demonstrated by multi-angle laser scattering (MALS) of Ctf4(CTD) and full-length Ctf4 (Fig. 1d and Extended Data Fig. 2) and non-denaturing nano-electrospray ionization mass spectrometry (native mass spectrometry) of Ctf4(CTD) (Fig. 1e).

We had previously shown that Ctf4 binds to the N-terminal portion of Pol1, the yeast orthologue of Pol  $\alpha$  (ref. 10). By progressive truncations of this largely unstructured region of Pol1, we identified a short linear motif spanning residues 137–149 that is necessary and sufficient for the association with Ctf4 *in vitro*. The motif has a mixed acidic and hydrophobic nature and is conserved from yeast to humans (Fig. 2b). Alanine scanning mutagenesis of the motif revealed that conserved residues F140, D142, I143, L144 and F147 are essential for the interaction with Ctf4 (Fig. 2c). The results of the biochemical experiments with recombinant proteins were confirmed by immunoprecipitation of Pol1 from extracts of yeast cells that were synchronized in G1 phase before release into S phase. Whereas wild-type Pol1 associated with Ctf4 and thus with the components of the Cdc45–MCM–GINS complex (Fig. 2d, control), the Pol1-A allele with alanine substitutions at D141, D142, L144 and F147 was unable to interact with either Ctf4 or the CMG (Fig. 2d, *pol1-A*).

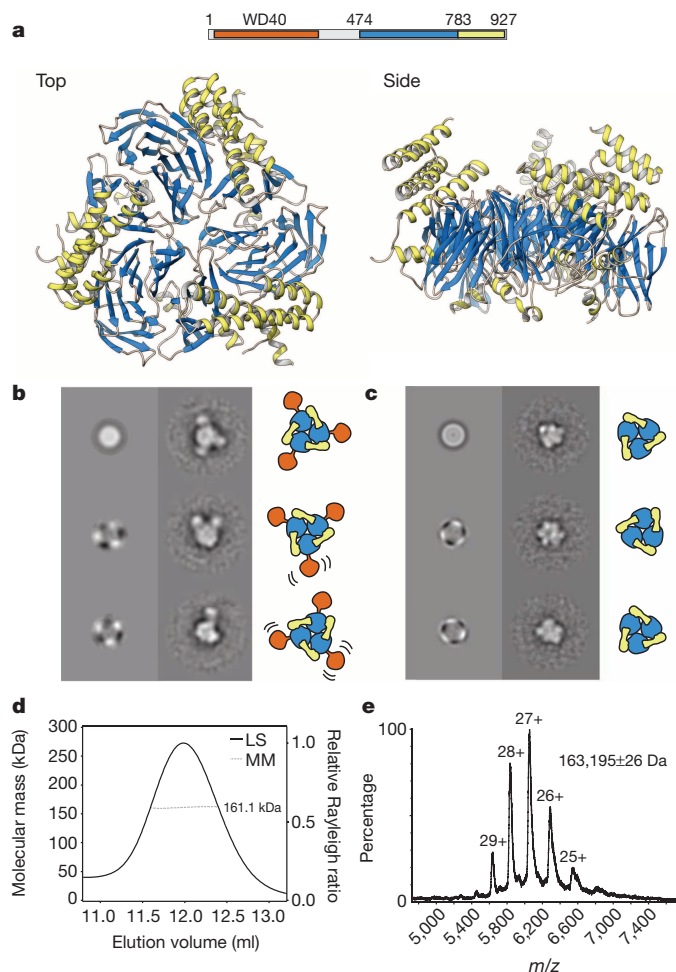
To define the structural basis for the interaction between Ctf4 and Pol  $\alpha$ , we soaked a 13-amino-acid peptide corresponding to Pol1 sequence 137-IDNFDDILGEFES-149 in the Ctf4(CTD) crystals. For the soaking experiments, we used a different crystal form of the Ctf4(CTD) trimer that is easier to grow; this form captures a topologically open conformation of the Ctf4(CTD) trimer resembling a cracked ring (Extended Data Fig. 3 and Supplementary Video 1). No differences are observed in protomer structure between closed and open forms of the Ctf4 trimer, with the exception of the helical domain located at the gap in the open form that becomes disordered in the electron density map. The potential functional significance of the open and closed forms of the Ctf4 trimer and their interconversion in solution is presently unclear.

The crystal structure of the Ctf4(CTD)–Pol  $\alpha$  complex shows that the helical domain protruding from the discoidal trimer is responsible for binding the polymerase (Fig. 2e). In the structure, amino acids 140–FDDILGEFES-149 of Pol1 fold into a two-turn  $\alpha$ -helix that packs in

<sup>1</sup>Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK. <sup>2</sup>Clare Hall Laboratories, Cancer Research UK London Research Institute, London EN6 3LD, UK. <sup>3</sup>Cancer Research UK Manchester Institute, University of Manchester, Manchester M20 4BX, UK. <sup>4</sup>MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK. <sup>5</sup>Protein purification, Cancer Research UK London Research Institute, London WC2A 3LY, UK. <sup>6</sup>Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK. <sup>†</sup>Present addresses: Imperial College, South Kensington, London SW7 2AZ, UK (R.L.P.); Cancer Research UK London Research Institute, London WC2A 3LY, UK (M.E.I.).

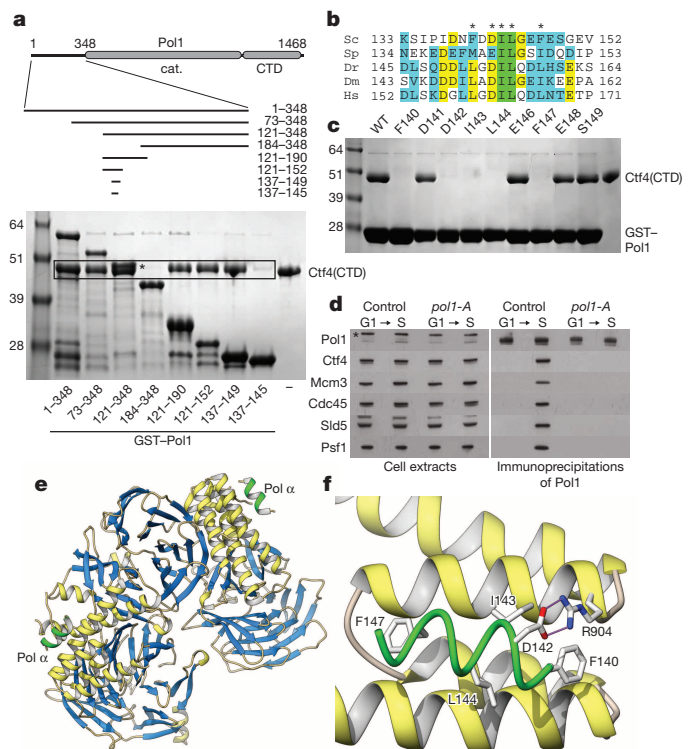
\*These authors contributed equally to this work.





**Figure 1 | Architecture of yeast Ctf4.** **a**, Ctf4 self-associates in a trimer of novel design. The panel shows top and side views of the crystal structure of the C-terminal region of yeast Ctf4 (Ctf4(CTD); amino acids 471–927). The protein is drawn as ribbon representation, coloured according to its domain structure: the  $\beta$ -propeller domain is in light blue and the helical domain in yellow. Above the drawing, a bar diagram shows the domain structure of full-length yeast Ctf4 and the extent of the region crystallized in our study. **b**, Analysis of full-length Ctf4 by single-particle electron microscopy. Multivariate statistical symmetry analysis detects a three-fold symmetry component for the full-length Ctf4 particle. Reference-free class averages of full-length Ctf4 reveal a core structure flexibly linked to up to three satellite domains. **c**, Analysis of Ctf4(CTD) by single-particle electron microscopy. The C-terminal domain of Ctf4 maintains a trimeric structure, as shown by multivariate statistical symmetry analysis and reference-free class averages. **d**, Size exclusion chromatography-multi-angle laser scattering analysis of yeast Ctf4(CTD). The light scattering is plotted alongside the fitted molecular masses. The protein eluted in a single peak, corresponding to a measured molecular mass of 161.1 kDa. The predicted molecular mass for the trimeric species is 163.1 kDa. **e**, Native mass-spectrometry analysis of yeast Ctf4(CTD). The measured molecular mass of 163,195 Da matches closely the predicted molecular mass of 163,148 Da for a trimeric species.

antiparallel fashion against helices  $\alpha 3$  and  $\alpha 5$ , on the outward-facing side of the helical domain of one Ctf4 protomer. The Ctf4-binding motif of Pol1 occupies each of the two binding sites available in this crystal form of the Ctf4(CTD) trimer. The interaction with the Pol1 peptide does not induce an appreciable conformational change in the helical domain of Ctf4(CTD) nor does it alter its position in the trimeric structure. The side chains of Pol1 residues F140, I143, L144 and F147, which were critical for the interaction in the pull-down assay, become buried at the interface and pack against a continuous hydrophobic surface formed by Ctf4 residues L867, A871, A894, A897 and I901 (Fig. 2f). The interaction is augmented



**Figure 2 | Pol  $\alpha$  contains a Ctf4-interacting motif that binds to the helical domain of Ctf4.** **a**, Identification of the Ctf4-binding motif of Pol1, the yeast orthologue of Pol  $\alpha$ . GST-tagged constructs spanning progressively smaller N-terminal regions of Pol1 were tested for interaction with Ctf4(CTD) in pull-down experiments on glutathione sepharose beads. The top panel shows the boundaries of the GST–Pol1 constructs; the bottom panel shows the result of the pull-down experiments, analysed by SDS–PAGE. The last lane on the right-hand side of the gel contains only Ctf4(CTD). The position of the Ctf4(CTD) band in the pull-down experiments is highlighted by a box. The asterisk marks the position of GST–Pol1(121–348), which overlaps partially with Ctf4(CTD). The first lane contains the protein size markers, with their molecular masses reported in kilodaltons on the left-hand side of the gel. **b**, Multiple sequence alignment of the Ctf4-binding motif of yeast Pol1 (Sc; *Saccharomyces cerevisiae*) with Pol  $\alpha$  sequences from *Schizosaccharomyces pombe* (Sp), *Danio rerio* (Dr), *Drosophila melanogaster* (Dm) and *Homo sapiens* (Hs). Invariant residues are highlighted in green, identical residues in yellow and similar residues in cyan. The asterisk marks amino acids that are essential for interaction with Ctf4(CTD) (see panel c). **c**, Alanine-scanning mutagenesis of the Ctf4-binding motif. Pol1 residues 137–149 were fused to GST and each amino acid between 140 and 149 (except G145) was mutated to alanine. The effect of each single-point mutation on the interaction with Ctf4(CTD) was tested by GST pull-down and analysed by SDS–PAGE. WT, wild type. The first lane contains the protein size markers, with their molecular masses reported in kilodaltons on the left-hand side of the gel. **d**, The budding yeast strains *POL1-9MYC* (control) and *pol1-A-9MYC* (*pol1-A*, containing the D141A, D142A, L144A and F147A mutations in the endogenous *POL1* locus) were grown at 24 °C, arrested in G1 phase and released into S phase for 30 min. The Myc-tagged proteins were isolated from cell extracts by immunoprecipitation on anti-Myc beads and the indicated proteins were detected by immunoblotting with the corresponding antibodies<sup>23</sup>. **e**, Co-crystal structure of Ctf4(CTD) bound to a peptide corresponding to the Ctf4-binding motif of Pol  $\alpha$ . Ctf4 is drawn as in Fig. 1a, the Ctf4-binding motif of Pol  $\alpha$  is drawn as a green ribbon. **f**, Detailed view of the interaction between the Ctf4-binding motif of Pol  $\alpha$  (green tube) and the helical domain of Ctf4 (yellow ribbon). The side chains of Pol1 residues F140, D142, I143, L144, F147 and Ctf4 residue R904 are shown as stick representation.

by polar contacts on the perimeter of the hydrophobic interface, between acidic residues D141, D142, E146 and E148 of Pol1 and basic residues K864, R868, R893, K900 and R904 of Ctf4 (Extended Data Figs 4a and 5). The salt bridge between the conserved residues D142 of Pol1 and R904

in helix  $\alpha 5$  of Ctf4 seems to be particularly important for binding, as the D142A mutation abolishes the interaction (Fig. 2c).

We next set out to investigate the mode of Ctf4 interaction with the CMG helicase. The association between Ctf4 and GINS is sufficiently strong to be assayed by size-exclusion chromatography<sup>10</sup> (Fig. 3a). We found that the interaction is dependent on the unstructured tail at the N terminus of the Sld5 subunit of GINS (Fig. 3a). Sequence comparison of fungal Sld5 orthologues revealed a conserved pattern of amino acids that is highly similar to the Ctf4-binding motif of Pol1 (Fig. 3b). Further dissection of the N-terminal tail of Sld5 confirmed that the initial

18 residues, containing the Ctf4-binding motif, were essential for the interaction (Fig. 3c). Our identification of a Ctf4-binding site in Sld5 is in agreement with earlier reports<sup>9,10</sup>, which had identified Sld5 as a Ctf4-binding subunit of GINS. Whether a conserved Ctf4-binding motif is present in the Sld5 sequence of higher eukaryotes is presently unclear. Alanine scanning mutagenesis of the Ctf4-binding motif of Sld5 highlighted the importance of the same pattern of conserved hydrophobic residues that were essential for Pol1 binding to Ctf4. Alanine mutation of I5, I8 and L9, corresponding to residues F140, I143 and L144 of Pol1, abolished the association of Sld5 with Ctf4, whereas mutation of L12, equivalent to Pol1 F147, weakened the interaction (Fig. 3d).

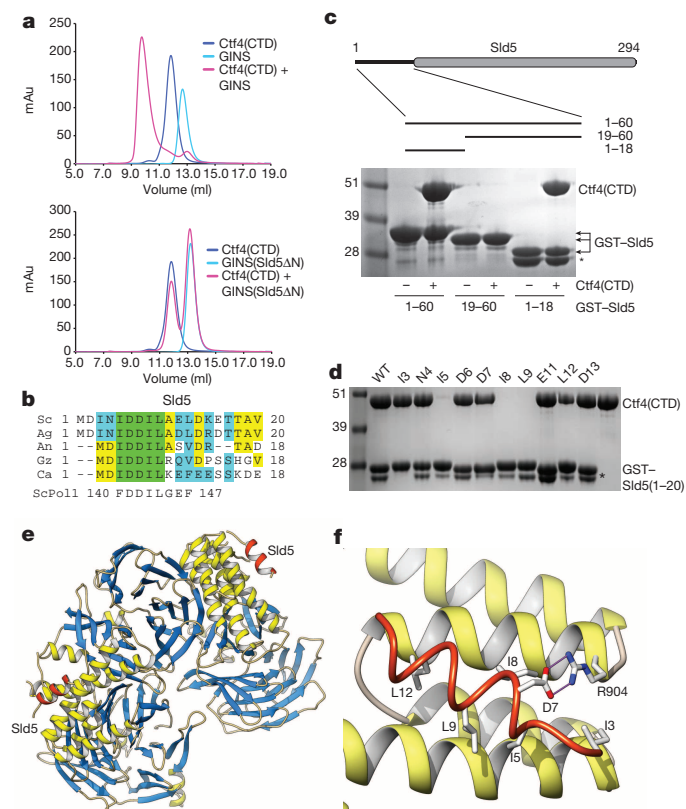
To ascertain whether the similarity in the two Ctf4-binding motifs extended to their mechanism of interaction, we soaked the Ctf4(CTD) crystals with the peptide MDINIDDILAE LDKETTAV, corresponding to amino acids 1–19 of yeast Sld5. The crystallographic analysis revealed a near-identical mode of interaction between Sld5 and Ctf4 as previously observed for the Ctf4-binding motif of Pol1 (Figs 3e and 4a and Extended Data Figs 4b and 5). The Ctf4-binding sequence of Sld5 includes an additional hydrophobic contact of I3 with Ctf4 residues A871, C874 and I901, which might account for its tighter (~5-fold) association with Ctf4 compared to Pol1 (Fig. 3f and Extended Data Fig. 6a). Conversely, the contribution of polar contacts seems diminished, as disruption of the salt bridge between D7 in Sld5 and R904 in Ctf4 does not impair appreciably the interaction (Fig. 3d, f). Collectively, these findings establish that Pol1 and GINS contain a Ctf4-binding motif that is conserved in sequence and function.

In yeast cells, Ctf4 appears to associate more tightly with the CMG helicase than with Pol  $\alpha$ , as the association of Ctf4 with CMG resists buffers containing 700 mM salt, whereas the Ctf4-dependent association of Pol  $\alpha$  with the replisome is lost at 300 mM salt<sup>23</sup>. Consistent with this, Ctf4 remains associated with CMG in cells containing mutations that disrupt the Ctf4-binding site of Sld5 (Extended Data Fig. 6b). It thus seems likely that Ctf4 has a more complex interaction with the CMG than with Pol  $\alpha$  and that additional contacts between Ctf4 and CMG remain to be characterized.

Our findings predict that Ctf4 can support simultaneous interactions of varying stoichiometry and with multiple partners (Fig. 4b). Indeed, native mass-spectrometry analysis of Ctf4(CTD) in the presence of the Ctf4-binding sequences of Pol  $\alpha$  and Sld5 showed reconstitution of complexes with 1:1, 1:2 and 1:3 Ctf4-to-peptide stoichiometries (Fig. 4c). To determine whether a Ctf4 trimer could support concomitant binding of three partner molecules, we analysed by EM reconstituted Ctf4(CTD)–GINS complexes (Extended Data Fig. 7). In agreement with the mass-spectrometry data, the EM analysis demonstrated the presence of 1, 2 or 3 copies of GINS bound to one Ctf4(CTD) trimer (Fig. 4d and Supplementary Video 2), each arranged radially around the Ctf4(CTD) trimer. Interestingly, each GINS molecule occupies a fixed position relative to the Ctf4 trimer, indicating that the interface between Ctf4 and GINS extends beyond the contact provided by the flexible N-terminal tail of Sld5. The Ctf4–GINS interface discernible in our EM averages is probably important to sustain the association between Ctf4 and the CMG helicase in the replisome.

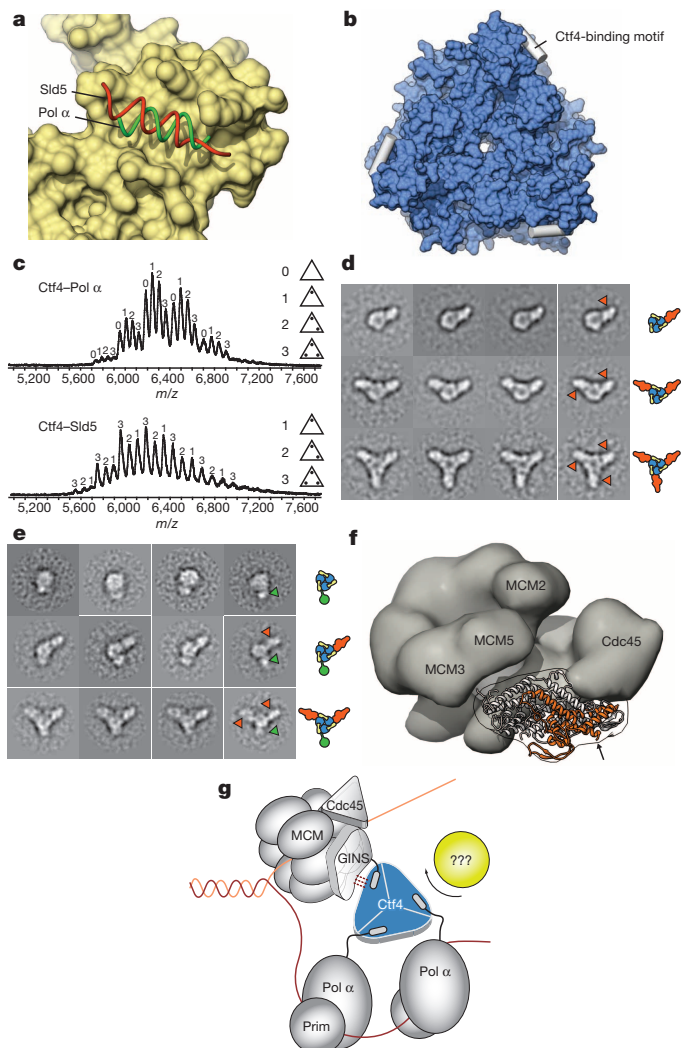
The reported function of Ctf4 as a physical link between helicase and polymerase prompted us to determine whether GINS and Pol  $\alpha$  can simultaneously associate with the Ctf4 trimer. We visualized by EM reconstituted hetero-assemblies of Ctf4(CTD) bound concurrently to GINS and the N-terminal region of Pol1 (residues 1–351) fused to protein A (Pol1(NTD)) (Extended Data Figs 8 and 9). As predicted by the trimeric nature of Ctf4, we could detect Ctf4(CTD)–GINS–Pol1(NTD) complexes of varied stoichiometries, with partial or full occupancy of the Ctf4 trimer (Fig. 4e). These data establish a structural basis for Ctf4 as the bridging factor between the CMG helicase and DNA polymerase  $\alpha$  in eukaryotic replication.

Inside the cell, the appropriate stoichiometry will presumably be determined by the constraints imposed on replisome assembly during replication initiation. Within the replisome, one binding site of the Ctf4 trimer



**Figure 3 | The Sld5 subunit of yeast GINS shares a common mechanism of Ctf4 binding with Pol  $\alpha$ .** **a**, Analysis of the Ctf4–GINS interaction by gel filtration chromatography, using Ctf4(CTD) and versions of GINS that contain either full-length (top panel) or N-terminally truncated Sld5 (Sld5ΔN; bottom panel). **b**, Multiple sequence alignment of the N terminus of fungal Sld5 sequences (Sc, *Saccharomyces cerevisiae*; Ag, *Ashbya gossypii*; An, *Aspergillus niger*; Gz, *Gibberella zeae*; Ca, *Candida albicans*). Invariant residues are highlighted in green, identical residues in yellow and similar residues in cyan. The Ctf4-binding motif of yeast Pol1 is reported below the alignment. **c**, Mapping of the Ctf4-binding sequence in the N terminus of Sld5 by GST pull-down analysis. The top panel shows the boundaries of the GST–Sld5 constructs tested for interaction with Ctf4(CTD); the bottom panel shows the results of the pull-down experiments, analysed by SDS–PAGE. The band marked with an asterisk corresponds to free GST. The first lane contains the protein size markers, with their molecular masses reported in kilodaltons on the left-hand side of the gel. **d**, Alanine-scanning mutagenesis of the Ctf4-binding motif. Residues 1–20 of yeast Sld5 were fused to GST and each position between 3 and 13 (except A10) was mutated to alanine. The effect of each single-point mutation on the interaction with Ctf4(CTD) was tested by GST pull-down and analysed by SDS–PAGE. WT, wild type. The first lane contains the protein size markers, with their molecular masses reported in kilodaltons on the left-hand side of the gel. **e**, Co-crystal structure of Ctf4(CTD) bound to a peptide corresponding to the Ctf4-binding motif of Sld5. Ctf4 is drawn as in Fig. 1a, the Ctf4-binding motif of Sld5 is drawn as red ribbon. **f**, Detailed view of the interaction between the Ctf4-binding motif of Sld5 (red tube) and the helical domain of Ctf4 (yellow ribbon). The side chains of Sld5 residues I3, I5, D7, I8, L9, L12 and Ctf4 residue R904 are shown in stick representation.





**Figure 4 | The Ctf4 trimer coordinates the recruitment of replication factors to the fork.** **a**, Superposition of the structures of Ctf4(CTD) bound to the Ctf4-binding motif of Pol  $\alpha$  (green tube) or Sld5 (red tube). Ctf4(CTD) is displayed as molecular surface, in light brown. **b**, Ctf4(CTD) can associate in principle with up to three partner proteins. To illustrate this point, the Ctf4-binding motif of Pol  $\alpha$  was modelled in each of the three binding sites of Ctf4(CTD). The helical Ctf4-binding motif is shown as a white cylinder, and Ctf4(CTD) is drawn as a molecular surface, in light blue. **c**, Native mass spectrometry analysis of the Ctf4(CTD) trimer in the presence of peptides corresponding to the Ctf4-binding motifs of Pol  $\alpha$  (top) and Sld5 (bottom). **d**, Single-particle electron microscopy analysis of the interaction of GINS with the Ctf4(CTD) trimer. Reference-free class averages of Ctf4(CTD) bound to one (top row), two (middle row) or three copies (bottom row) of GINS are shown. **e**, Reference-free class averages of the Ctf4(CTD)-Pol1(NTD) (top row), Ctf4(CTD)-Pol1(NTD)-GINS (middle row) and Ctf4(CTD)-Pol1(NTD)-(GINS)<sub>2</sub> (bottom row) heteroassemblies. **f**, The panel shows the crystal structure of human GINS (ref. 25) docked into the electron microscopy reconstruction of the CMG helicase (ref. 24). The Sld5 subunit of GINS is coloured orange and the rest of GINS is shown in white. The density for the MCM and Cdc45 subunits of the CMG helicase is shown as a grey surface, whereas the density of the GINS tetramer is shown as an outline. The position of MCM2, MCM3, MCM5 and Cdc45, which surround GINS in the helicase complex, is indicated. An arrow marks the N-terminal residue in the Sld5 structure. **g**, A model of Ctf4 function at the replication fork, as the physical bridge between the CMG helicase and the DNA polymerase  $\alpha$ /primase complex. The additional contacts between Ctf4 and GINS suggested by the EM analysis (panel **d**) are indicated by dashed lines.

is likely to engage in a constitutive interaction with GINS, to anchor Ctf4 to the CMG helicase at the fork. In the molecular model of the CMG<sup>24</sup>, the GINS structure<sup>25</sup> has the Sld5 N terminus favourably positioned for

binding Ctf4, in agreement with our biochemical findings (Fig. 4f). As replisomes formed at a replication origin need not remain physically tethered for efficient replication<sup>26</sup>, it is unlikely that Ctf4 acts by coupling two CMG helicases. The other two binding sites of the CMG-bound Ctf4 trimer would remain available for interaction with Pol  $\alpha$ , indicating that two copies of the Pol  $\alpha$ -primase complex might work together during lagging-strand synthesis (Fig. 4g). Such coupling of helicase and polymerase in the eukaryotic replisome would be functionally analogous to the emerging model of the *E. coli* replisome, where two DNA polymerases cooperate in lagging-strand synthesis to increase processivity and efficiency of nucleotide polymerization<sup>5–7</sup>.

In addition to its function as a helicase-polymerase bridge, Ctf4 seems to be ideally suited to fulfil a wider role in replication, as a platform for coordinating the activity of replication factors at the fork. In this model, one Ctf4 protomer would keep the trimer constitutively anchored to the CMG, whereas other replisome components, including Pol  $\alpha$ , would engage with the helicase in a dynamic interaction mediated by the Ctf4-binding motif identified here. We note that this model of Ctf4 function is reminiscent of the way the proliferating cell nuclear antigen (PCNA) interacts with replication factors such as Fen1 and DNA ligase I (ref. 27). Thus, in addition to bridging CMG helicase and Pol  $\alpha$ , Ctf4 might recruit to the fork other factors required for efficient replication under normal conditions or needed to deal with exceptional situations during replicative stress.

## METHODS SUMMARY

Included in the Methods section are full experimental procedures for the crystallographic and electron microscopy analysis of Ctf4 and its complexes with Pol  $\alpha$  and GINS, the biophysical characterization of Ctf4 by multi-angle laser scattering and native mass spectrometry, and the *in vitro* and *ex vivo* biochemical analysis of the interaction of Ctf4 with Pol  $\alpha$  and Sld5.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 October 2013; accepted 6 March 2014.

Published online 4 May 2014.

- DePamphilis, M. L. & Bell, S. D. *Genome Duplication* (Garland Science, 2011).
- Zegerman, P. & Diffley, J. F. X. DNA replication as a target of the DNA damage checkpoint. *DNA Repair (Amst.)* **8**, 1077–1088 (2009).
- Cortez, D. Unwind and slow down: checkpoint activation by helicase and polymerase uncoupling. *Genes Dev.* **19**, 1007–1012 (2005).
- Errico, A. & Costanzo, V. Mechanisms of replication fork protection: a safeguard for genome stability. *Crit. Rev. Biochem. Mol. Biol.* **47**, 222–235 (2012).
- Lia, G., Michel, B. N. D. & Allemand, J.-F. S. S. Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science* **335**, 328–331 (2012).
- Georgescu, R. E., Kurth, I. & O'Donnell, M. Single-molecule studies reveal the function of a third polymerase in the replisome. *Nature Struct. Mol. Biol.* **19**, 113–116 (2012).
- McInerney, P., Johnson, A., Katz, F. & O'Donnell, M. Characterization of a triple DNA polymerase replisome. *Mol. Cell* **27**, 527–538 (2007).
- Reyes-Lamothe, R., Sherratt, D. J. & Leake, M. C. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328**, 498–501 (2010).
- Tanaka, H. *et al.* Ctf4 coordinates the progression of helicase and DNA polymerase  $\alpha$ . *Genes Cells* **14**, 807–820 (2009).
- Gambus, A. *et al.* A key role for Ctf4 in coupling the MCM2–7 helicase to DNA polymerase  $\alpha$  within the eukaryotic replisome. *EMBO J.* **28**, 2992–3004 (2009).
- Sengupta, S. *et al.* Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.* **23**, 543–552 (2013).
- Lou, H. *et al.* Mrc1 and DNA polymerase  $\epsilon$  function together in linking DNA replication and the S phase checkpoint. *Mol. Cell* **32**, 106 (2008).
- Zhu, W. *et al.* Mcm10 and And-1/CTF4 recruit DNA polymerase  $\alpha$  to chromatin for initiation of DNA replication. *Genes Dev.* **21**, 2288–2299 (2007).
- Kouprina, N. *et al.* CTF4 (CHL15) mutants exhibit defective DNA metabolism in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**, 5736–5747 (1992).
- Miles, J. & Formosa, T. Evidence that POB1, a *Saccharomyces cerevisiae* protein that binds to DNA polymerase  $\alpha$ , acts in DNA metabolism *in vivo*. *Mol. Cell. Biol.* **12**, 5724–5735 (1992).
- Gosnell, J. A. & Christensen, T. W. *Drosophila* Ctf4 is essential for efficient DNA replication and normal cell cycle progression. *BMC Mol. Biol.* **12**, 13 (2011).
- Bermudez, V. P., Farina, A., Tappin, I. & Hurwitz, J. Influence of the human cohesion establishment factor Ctf4/AND-1 on DNA replication. *J. Biol. Chem.* **285**, 9493–9505 (2010).

18. Williams, D. & McIntosh, J. mcl1+, the *Schizosaccharomyces pombe* homologue of CTF4, is important for chromosome replication, cohesion, and segregation. *Eukaryot. Cell* **1**, 758–773 (2002).
19. Tanaka, H. *et al.* Replisome progression complex links DNA replication to sister chromatid cohesion in *Xenopus* egg extracts. *Genes Cells* **14**, 949–963 (2009).
20. Petronczki, M. *et al.* Sister-chromatid cohesion mediated by the alternative RF-CCtf18/Dcc1/Ctf8, the helicase Chl1 and the polymerase- $\alpha$ -associated protein Ctf4 is essential for chromatid disjunction during meiosis II. *J. Cell Sci.* **117**, 3547–3559 (2004).
21. Hanna, J. S., Kroll, E. S., Lundblad, V. & Spencer, F. A. *Saccharomyces cerevisiae* CTF18 and CTF4 are required for sister chromatid cohesion. *Mol. Cell. Biol.* **21**, 3144–3158 (2001).
22. Yoshizawa-Sugata, N. & Masai, H. Roles of human AND-1 in chromosome transactions in S phase. *J. Biol. Chem.* **284**, 20718–20728 (2009).
23. Gambus, A. *et al.* GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biol.* **8**, 358–366 (2006).
24. Costa, A. *et al.* The structural basis for MCM2–7 helicase activation by GINS and Cdc45. *Nature Struct. Mol. Biol.* **18**, 471–477 (2011).
25. Chang, Y. P., Wang, G., Bermudez, V., Hurwitz, J. & Chen, X. S. Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc. Natl Acad. Sci. USA* **104**, 12685–12690 (2007).
26. Yardimci, H., Loveland, A. B., Habuchi, S., van Oijen, A. M. & Walter, J. C. Uncoupling of sister replisomes during eukaryotic DNA replication. *Mol. Cell* **40**, 834–840 (2010).
27. Beattie, T. R. & Bell, S. D. Coordination of multiple enzyme activities by a single PCNA in archaeal Okazaki fragment maturation. *EMBO J.* **31**, 1556–1567 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We would like to thank L. Collinson and R. Carzaniga (LRI) for EM access; D. Frith and B. Snijder (LRI) for mass spectrometry work; J. Maman for help with SEC-MALS; and P. Zegerman and J. Gannon for comments on the manuscript. This work was supported by the Gates Cambridge PhD program (A.C.S.), CRUK (A.C. and K.L.), MRC (K.L.) and a Wellcome Trust SRF award in basic biomedical science (L.P.).

**Author Contributions** A.C. and L.P. conceived the project; A.C.S., J.C.Z., R.L.P., K.L., A.C. and L.P. designed experiments; A.C.S., J.C.Z., R.L.P., C.E., F.v.D., M.E.I., M.L.K., L.R., S.K. and D.M.-V. performed experiments; A.C.S., J.C.Z., D.M.-V., K.L., A.C. and L.P. analysed the data; A.C. and L.P. wrote the paper, with contributions and critical comments from the other authors.

**Author Information** Coordinates and structure factors for Ctf4(CTD) (selenomethionine-labelled protein), Ctf4(CTD) (native), Ctf4(CTD)-Pol  $\alpha$  and Ctf4(CTD)-Sld5 complexes are available from the Protein Data Bank under accession codes 4C8H, 4C8S, 4C93 and 4C95, respectively. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.P. (lp212@cam.ac.uk) or A.C. (alessandro.costa@cancer.org.uk).



## METHODS

**DNA constructs for X-ray crystallography, MALS and MS of Ctf4(CTD) and biochemical analysis of the Ctf4(CTD)–PolI and Ctf4(CTD)–Sld5 interactions.** Fold recognition analysis in Phyre2 (ref. 28) predicted that the C-terminal half of yeast Ctf4, responsible for interactions with GINS and Pol  $\alpha$ , contained a WD40 domain fused to an  $\alpha$ -helical region. A region of yeast Ctf4 comprising amino acids 471–927 (natural C-end; Ctf4(CTD)) was PCR amplified from *S. cerevisiae* genomic DNA and cloned into a bacterial pRSFDuet-1 T7 expression plasmid (Novagen) via unique BamHI and AvrII sites. Using PCR primer extension, a TEV protease site was introduced at the start of the Ctf4(CTD) open reading frame sequence and after the N-terminal His<sub>6</sub>-affinity tag encoded by the pRSFDuet-1 vector.

The DNA polymerase  $\alpha$  (Pol  $\alpha$ )- and Sld5–GST fusion constructs used in pull-down experiments were generated by insertion of the appropriate nucleotide sequence into the NcoI and XhoI sites of the pGAT2 T7 expression plasmid encoding a thrombin-cleavable N-terminal GST fusion protein<sup>29</sup>.

A construct for bacterial expression of yeast GINS was prepared starting from vector pKL653 (ref. 10), by subcloning one expression cassette comprising *psf3* and *psf1AC* (amino acids 1–164) into the NcoI and NotI sites in the first MCS of a pRSFDuet-1 expression plasmid, and another expression cassette comprising *psf2* with an N-terminal His<sub>6</sub> affinity tag and *sld5* into the second MCS of pRSFDuet-1, resulting in the polycistronic pGINS–Duet-1 expression plasmid. The GINS(Sld5AN) construct used for analytical gel filtration experiments was derived from the pGINS–Duet-1 vector, by replacing the second expression cassette with a modified cassette that encodes, in addition to His<sub>6</sub>–*psf2*, a version of *sld5* coding for a truncated protein lacking the first 48 amino acids at its N terminus.

**DNA constructs for electron microscopy and MALS of full-length Ctf4.** Full-length *S. cerevisiae* Ctf4 and Ctf4 N-terminal deletion (Ctf4(CTD), amino acids 461–927) constructs were both cloned into the pET28c vector (Novagen) to express an N-terminal His<sub>6</sub> affinity tag. The *S. cerevisiae* GINS Psf1 C-terminal deletion ( $\Delta$ CT, amino acids 1–164) construct was subcloned from a previously described GINS operon-containing plasmid<sup>10</sup> into the pET28c vector and carries an N-terminal Strep III tag in the GINS Psf3 subunit. The PolI–protein A fusion was subcloned into the pET Strep II–TEV LIC vector (QB3 MacroLab) by ligation independent cloning<sup>30</sup>. This construct contains in the following order: an N-terminal Strep II tag, the N-terminal domain (amino acids 1–351) of *S. cerevisiae* PolI, the protein A region of the TAP tag<sup>31</sup> and a C-terminal His<sub>7</sub> affinity tag.

**Protein expression and purification for X-ray crystallography, MALS and MS of the Ctf4(CTD) and biochemical analysis of the Ctf4(CTD)–PolI and Ctf4(CTD)–Sld5 interactions.** Ctf4(CTD) was overexpressed in *E. coli* strain BL21 (DE3)Rosetta2 with IPTG induction and overnight expression at 20 °C in LB medium. After overexpression, 4 l of cells were collected and re-suspended in 50 mM Tris pH 7.0, 500 mM NaCl, 10% (w/v) glycerol, 1 mM DTT and protease inhibitors (Sigma). Cells were lysed via sonication, the crude extract was clarified by centrifugation and the supernatant was applied to a 4-ml column of nickel agarose resin (Sigma) using gravity flow. The column with bound Ctf4(CTD) was washed in buffer supplemented with 20 mM imidazole and Ctf4(CTD) elution was performed with buffer supplemented with 200 mM imidazole. Eluted Ctf4(CTD) was further purified by gel filtration chromatography over a Superdex 200 16/60 HiLoad column (GE Healthcare) in 25 mM HEPES pH 7.0, 200 mM NaCl and 10% (w/v) glycerol and peak fractions were pooled, concentrated to 10 mg ml<sup>−1</sup>, flash frozen in liquid nitrogen and stored in small aliquots at −80 °C. Selenomethionine labelling of Ctf4(CTD) was achieved by metabolic inhibition of the methionine pathway<sup>32</sup> and overnight expression as for the wild-type protein. The selenomethionine-labelled protein was purified in the same way as the native Ctf4(CTD) except that all buffers were supplemented with 10 mM DTT.

GINS constructs were overexpressed in *E. coli* strain BL21 (DE3)Rosetta2 with IPTG induction and overnight expression at 25 °C in LB medium. After overexpression, 8 l of cells were harvested and re-suspended in 50 mM Tris pH 7.0, 500 mM NaCl, 10% (w/v) glycerol, 1 mM DTT and protease inhibitor. Cells were lysed via sonication, the crude extract was clarified by centrifugation and the supernatant was applied to a 3-ml column of nickel agarose resin (Sigma) using gravity flow. GINS bound to beads was eluted with buffer supplemented with 10 mM imidazole. The salt concentration of the eluted GINS sample was adjusted to below 160 mM NaCl and the protein was applied to an ion-exchange 6-ml Resource Q column pre-equilibrated in 20 mM HEPES pH 8.0, 160 mM NaCl and eluted with a buffer gradient of 0.16 M to 0.5 M NaCl over 40 column volumes. Peak fractions containing GINS were pooled and further purified by gel filtration over a Superdex 200 16/60 column as described for Ctf4(CTD). Purified GINS samples were flash-frozen in liquid nitrogen and stored in small aliquots at −80 °C.

**Protein expression and purification for electron microscopy and MALS.** Each expression construct was transformed into BL21 (DE3)–CodonPlus cells (Stratagene) and 2 to 4 l of cells were grown to an optical density of 0.5 before induction with 1 mM IPTG, at 37 °C for 2 h. Each 2-l cell pellet was re-suspended in 40 ml lysis

buffer and cells were lysed via sonication. The resulting lysate was subject to the following purification steps.

**Ctf4 purification:** cleared lysate containing His-tagged full-length Ctf4 or Ctf4(CTD) was incubated with 1 ml Ni-NTA resin (Qiagen), washed with 20 ml buffer A (50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8.0, 300 mM NaCl, 20 mM imidazole) and eluted five times each with 1 ml buffer A containing 250 mM imidazole. The resulting elution was dialysed in 2 l 100 mM NaCl, 20 mM Tris pH 8.0, 1 mM DTT for 2 h with fresh buffer exchanged after the first hour. The dialysed elution was further purified by Mono Q (GE Healthcare) ion exchange through a 0.1–1 M NaCl gradient in 20 mM Tris pH 8.0, 1 mM DTT over 40 ml with 0.5 ml elutions. Peak fractions from the Mono Q were concentrated and polished via Superdex 200 16/600 HiLoad or 10/300 GL (GE Healthcare) size exclusion chromatography in buffer B (150 mM NaCl, 20 mM Tris pH 8.0). Peak elutions were pooled and concentrated to 5 mg ml<sup>−1</sup> and stored at −80 °C in 2 nmol aliquots.

**GINS purifications:** cleared lysate containing Strep III tagged GINS(Psf1AC) was incubated with 1 ml StrepTactin resin (IBA Life Sciences), washed with 20 ml buffer C (150 mM NaCl, 100 mM Tris pH 8.0) and eluted five times each with 1 ml buffer C supplemented with 2.5 mM desthiobiotin (IBA Life Sciences). Proteins were stored at −80 °C with a concentration of 1 mg ml<sup>−1</sup> in 2 nmol aliquots.

**PolI purification:** cleared lysate containing N-terminal Strep II- and C-terminal His-tagged PolI–NTD–Protein A (hereafter referred to as PolI (NTD)) was purified first via Ni-NTA, followed by StrepTactin affinity using the same method described above.

The typical yield from 2 l cells for full-length Ctf4, Ctf4(CTD), GINS(Psf1AC) and PolI (NTD) are around 0.15 mg, 0.65 mg, 2.5 mg and 1.5 mg, respectively. The identity of all proteins was confirmed by trypsinization/mass spectrometry using a LTQ OrbitrapXL instrument (Protein Analysis and Proteomics, LRI).

**In vitro reconstitution of recombinant protein complexes for electron microscopy.** For the Ctf4–GINS complex, 2 nmol of recombinant yeast Ctf4 (full-length or CTD) and GINS(Psf1AC) were co-incubated in 500 mM sodium acetate for 10 min on ice with a reaction volume of around 200  $\mu$ l. To achieve high reproducibility the following procedure was followed. The reconstitution mix was initially dialysed in 500 mM sodium acetate, 25 mM HEPES pH 7.6, 0.5 mM DTT for 1 h at 4 °C in dialysis tubes with 6,000–8,000 Da MWCO (GeBaflex). The dialysis buffer was changed hourly to contain progressively 400 mM, 300 mM, 200 mM, 150 mM sodium acetate. 100–150  $\mu$ l of the final reconstituted complex was separated via glycerol gradient sedimentation. For the Ctf4(CTD)–PolI (NTD) complex and the Ctf4(CTD)–GINS(Psf1AC)–PolI (NTD) complex, 2 nmol of recombinant Ctf4(CTD) was used, 1:2 and 1:1:1 mol ratios were applied respectively and the dialysis was performed in buffers containing progressively 400 mM, 300 mM, 200 mM, 100 mM, 50 mM sodium acetate.

**Glycerol gradient sedimentation with GraFix.** Glycerol gradient sedimentation of full-length Ctf4, Ctf4(CTD) and complexes of Ctf4(CTD) with GINS(Psf1AC) and PolI (NTD) was performed based on the GraFix method<sup>33</sup>. Briefly, 5 ml 10% to 30% or 15% to 35% glycerol gradients were poured either with or without 0% to 0.1% glutaraldehyde gradient. The protein or reconstituted protein complex was loaded on top of the gradient and centrifuged at 50,000 r.p.m., 4 °C in a SW 55 Ti ultracentrifuge rotor (Beckman Coulter) for 16 h. Fractions were collected manually from the top of the gradient, resolved through a 4%–12% polyacrylamide–gradient gel (Biorad) in MOPS buffer at room temperature and silver stained for analysis.

**Crystallization and structure determination of Ctf4(CTD).** Ctf4(CTD) crystals were grown by vapour diffusion in hanging drop, mixing equal volumes of Ctf4(CTD) protein at 10 mg ml<sup>−1</sup> and 0.2 M tri-sodium citrate pH 6.2, 7–9% PEG 8000 and 0.45–0.9 M NaCl at 19 °C. Ctf4(CTD) crystals appeared within 2–3 days and grew to full size over the course of 2 weeks. For structure determination, selenomethionine-labelled Ctf4(CTD) crystals were grown against 0.2 M tri-sodium citrate pH 6.2 and 8–10% PEG 8000 at 19 °C, using the same protein concentration and drop ratio as for the native protein.

X-ray diffraction data for selenomethionine-labelled Ctf4(CTD) crystals were collected at the peak wavelength of the selenium K-edge ( $\lambda = 0.97938$  nm) at beamline I03 of the Diamond Light Source. The data were integrated with XDS<sup>34</sup>, space group symmetry was assigned in POINTLESS and intensities scaled in AIMLESS<sup>35</sup>. The selenomethionine protein crystallized in the orthorhombic space group P2<sub>2</sub>2<sub>1</sub> with unit cell dimensions of  $a = 107.1$  Å,  $b = 118.1$  Å,  $c = 155.7$  Å and one Ctf4(CTD) trimer per asymmetric unit. The position of the selenium atoms was determined using the single-wavelength anomalous dispersion (SAD) method in PHENIX Autosol, an interpretable electron density map was calculated to a resolution of 2.7 Å and an initial model was generated using the PHENIX AutoBuild function<sup>36</sup>. The crystallographic model was extended and completed by repeated cycles of manual building in Coot and crystallographic refinement with PHENIX Refine<sup>36,37</sup>. The final model was refined using data to 2.7 Å, to R-work and R-free values of 0.1895 and 0.2284 and a Molprobity score of 1.15 (ref. 38). The following amino acids were not included in the final model due to missing or poor electron density and are

presumed to be disordered: 471 to 473, 644 to 647, 797 to 813 and 926 to 927 in chain A; 471 to 473, 644 to 647, 664 to 670, 794 to 813 and 924 to 927 in chain B; 471 to 473, 664 to 670, 794 to 813 and 926 to 927 in chain C. Statistics of data processing and crystallographic refinement are reported in Extended Data Table 1.

X-ray diffraction data for the native Ctf4(CTD) crystals were collected at beamline I04 of the Diamond Light Source and the data were processed as for the selenomethionine data set. The native protein crystallized in the same orthorhombic space group  $P2_12_1$  as the selenomethionine protein crystals, but with different unit cell dimensions  $a = 88.9 \text{ \AA}$ ,  $b = 100.0 \text{ \AA}$ ,  $c = 219.3 \text{ \AA}$ , caused by an alternative set of crystal contacts made by the Ctf4(CTD) trimer in the asymmetric unit. The structure of native Ctf4(CTD) was solved by molecular replacement in PHASER<sup>39</sup>, using the structure of one protomer of the selenomethionine Ctf4(CTD) trimer as search model. The final model was refined using data to 3.0 Å resolution, to R-work and R-free values of 0.1674 and 0.2049 and a Molprobity score of 1.42. The following amino acids were not included in the final model due to missing or poor electron density and are presumed to be disordered: 471 to 473, 664 to 670, 792 to 813 in chain A; 471 to 473, 797 to 813 in chain B; 471 to 473, 664 to 670, 777 to 927 (helical domain) in chain C. In this crystal form, the Ctf4(CTD) structure adopts a more open conformation where one interface between Ctf4 protomers widens to become a narrow gap and the helical domain of one of the Ctf4(CTD) protomers at the interface becomes disordered (Extended Data Fig. 4). Statistics of data processing and crystallographic refinement are reported in Extended Data Table 1.

**Co-crystallization of Ctf4(CTD) with Pol  $\alpha$  and Sld5.** For co-crystallization experiments, the peptides IDNFDDILGEFES and MDINIDDILAELDKETTAV, corresponding to amino acids 137–149 of yeast Pol1 and 1–19 of yeast Sld5, respectively, were synthesized. The Pol1 peptide was solubilized in the same buffer as purified Ctf4(CTD) to a concentration of 340  $\mu\text{M}$ ; the Sld5 peptide was solubilized in water to a concentration of 2 mM. Soaking was performed by adding 1  $\mu\text{l}$  of Pol1 peptide or 0.5  $\mu\text{l}$  of Sld5 peptide to a 2  $\mu\text{l}$  crystallization drop containing native Ctf4(CTD) crystals. The crystals were soaked with the peptide for 24 h at 19 °C, back-soaked in crystallization buffer and flash-frozen in liquid nitrogen. X-ray diffraction data for Ctf4(CTD) crystals soaked with the Pol  $\alpha$  and Sld5 peptides were collected on beamline I04 of the Diamond Light Source and processed as for the native crystals. The position of the Ctf4-binding motifs of Pol  $\alpha$  and Sld5 in the crystals structure of Ctf4(CTD) was readily identified by inspection of  $F_o - F_c$  difference Fourier maps. Amino acids 140–149 of Pol1 and 3–15 of Sld5 were built in the electron density map and the structures of Ctf4(CTD) bound to Pol  $\alpha$  and Sld5 were then further refined using Coot and PHENIX Refine to R-work/R-free values of 0.1718/0.2099 and 0.1787/0.2141, respectively. Molprobity scores for the Ctf4(CTD)–Pol  $\alpha$  and Ctf4(CTD)–Sld5 structures were 1.33 and 1.32, respectively. Statistics of data processing and crystallographic refinement are reported in Extended Data Table 1.

**Sample preparation for EM.** Negative stain analysis was performed using 400 mesh carbon coated grids (Agar Scientific). Carbon was evaporated onto freshly cleaved mica with a Q150TE coater (Quorum Technologies) and incubated overnight before floating. Dried carbon grids were glow discharged for 30–60 s at 45 mA using a 100 $\times$  glow discharger (Electron Microscopy Sciences). A 4- $\mu\text{l}$  drop of the peak fraction from each GraFix-processed sample was applied onto the grid. Subsequently, grids were sequentially laid on top of five distinct 75  $\mu\text{l}$  drops of 2% uranyl formate solution, and stirred for 10 s each time, before blotting to dryness.

**EM data collection.** Negative stain analyses of all complexes were performed using a Tecnai LaB6 G2 Spirit transmission electron microscope (FEI) operating at 120 keV (Electron Microscopy Unit, London Research Institute). Images were recorded using a 2k  $\times$  2k GATAN Ultrascan 100 camera at a nominal magnification of 30,000 $\times$  (3.45 Å/pixel at the specimen level). Between 100 and 350 micrographs were collected for each data set.

**Single-particle analysis.** CTF corrected image stacks were prepared in the EMAN2 environment<sup>40</sup>. Single-particle symmetry analysis was performed as described<sup>41</sup>. Reference-free two-dimensional class averages were calculated using the one-step rotation and classification approach as described<sup>24</sup>, followed by routine MSA/MRA IMAGIC protocols<sup>42</sup>.

**GST pull downs.** For each Pol  $\alpha$  and Sld5 construct to be tested for interaction with Ctf4(CTD), a 25-ml *E. coli* BL21 (DE3) culture overexpressing the GST fusion construct was pelleted, re-suspended in buffer 50 mM Tris pH 7.0, 500 mM NaCl, 10% (w/v) glycerol, 1 mM DTT and protease inhibitors (Sigma) and lysed by sonication. After centrifugation, the soluble extract was mixed with 50  $\mu\text{l}$  of Glutathione Sepharose beads (GE Healthcare) pre-equilibrated in the same buffer and incubated under rotation at 4 °C for 1 h. Unbound protein was removed by three consecutive washes with 1 ml of buffer, followed by 3 1-ml washes with pull-down buffer (20 mM HEPES pH 7.2, 150 mM NaCl, 5% (w/v) glycerol, 0.1% Igepal CA-630, 1 mM TCEP and 1% BSA). Subsequently, 500  $\mu\text{l}$  of purified Ctf4(CTD) protein at a concentration of 2 mg ml<sup>-1</sup> was added to the Sepharose beads and binding was allowed to take place for an additional hour at 4 °C. The binding reaction was stopped by two consecutive washes with 1 ml of pull-down buffer and a final 1-ml

wash with pull-down buffer without BSA. The Sepharose beads were mixed with SDS loading dye and Ctf4(CTD) interactions with the respective bait proteins were detected via SDS–PAGE. As a control, Ctf4(CTD) was tested for unspecific interaction with the Glutathione Sepharose resin and with GST and in both cases no interaction was detected.

**Yeast strains and growth.** The yeast strains CC2619 (*POL1-9MYC pep4 $\Delta$ ::URA3 ADE2*) and CC10682 (*pol1-A POL1-9MYC pep4 $\Delta$ ::URA3 ADE2*) were grown at 24 °C in rich medium (1% yeast extract, 2% peptone, 40 g/ml adenine) with 2% glucose as carbon source. Cells were synchronized in G1 phase by adding 7.5  $\mu\text{g ml}^{-1}$  alpha factor mating pheromone for 70% of one generation time, followed by additional aliquots of 2.5  $\mu\text{g ml}^{-1}$  every 20 min up to 1.5 generation times.

**Immunoprecipitation of proteins from yeast cell extracts.** Myc-tagged proteins were isolated from yeast cell extracts as described previously<sup>43</sup>.

**Gel filtration.** All proteins were purified as described above. Putative complexes were reconstituted before analytical gel filtration by mixing stoichiometric molar ratios of Ctf4(CTD) with either GINS(Psf1 $\Delta$ C) or GINS(Psf1 $\Delta$ C,Sld5 $\Delta$ N) followed by centrifugation at 16,000g for 10 min at 4 °C to remove potential aggregates. 100  $\mu\text{l}$  samples of Ctf4(CTD), GINS(Psf1 $\Delta$ C), GINS(Psf1 $\Delta$ C,Sld5 $\Delta$ N), Ctf4(CTD)–GINS (Psf1 $\Delta$ C) and Ctf4(CTD)–GINS(Psf1 $\Delta$ C,Sld5 $\Delta$ N) were subsequently fractionated over a Superdex S200 HR 10/300 column (GE Healthcare) pre-equilibrated in 20 mM HEPES pH 7.2, 140 mM KCl.

**MALS analysis of Ctf4(CTD).** 100  $\mu\text{l}$  of Ctf4(CTD) protein at a concentration of 2 mg ml<sup>-1</sup> was loaded onto a Superdex S200 HR 10/300 gel-filtration column (GE Healthcare) in 25 mM HEPES pH 7.0, 200 mM NaCl at a flow rate of 0.5 ml min<sup>-1</sup>. The column was controlled using an Äkta Purifier System (GE Healthcare) and was linked to a DAWN 8+ 8-angle light scattering detector (Wyatt Technology) with a fused silica sample cell using a laser wavelength of 664 nm. The change in the refractive index was detected using an Optilab T-REX refractometer with extended range (Wyatt Technology) using a wavelength of 658 nm. Data collection and analysis was carried out using the ASTRA6 software package (Wyatt Technology). Molecular mass determination across the sample peak was carried out using a Zimm-plot derived global fitting algorithm with a fit degree of 1 and a dn/dc value of 0.1850 ml g<sup>-1</sup>.

**MALS analysis of full-length Ctf4.** Around 100  $\mu\text{g}$  of Ctf4 full-length protein was loaded onto a Wyatt MP-030S5 HPLC size-exclusion chromatography column (Wyatt) mounted on an AKTA Micro (GE Healthcare) chromatography. The column was equilibrated in a buffer containing 150 mM NaCl, 20 mM Tris pH 8.0, 1 mM DTT. The chromatography system was coupled to an 8-angle light scattering detector (DAWN 8+) and a refractive index detector (OptiLab T-Rex) (Wyatt Technology). Data were collected every 0.5 s. Data analysis was carried out using ASTRA VI.

**Fluorescence polarization.** Both the Pol1 peptide (137-IDNFDDILGEFES-149) and the yeast Sld5 peptide (1-MDINIDDILAELDKETTAV-19) were synthesized with an N-terminal fluorescein label. The lowest concentration of peptide at which the binding studies could be performed was determined via peptide calibration curves. Fluorescence anisotropy measurements were recorded in a PHERAstar Plus multi-detection plate reader (BMG Labtech) equipped with fluorescence polarization optic module ( $\lambda_{\text{ex}} = 485 \text{ nm}$ ;  $\lambda_{\text{em}} = 520 \text{ nm}$ ) at 25 °C. Each data point is the mean of 200 flashes per well. The voltage gain was set by adjusting the target mP values of fluorescein-labelled peptides relative to that of fluorescein (35 mP). Serial dilutions of Ctf4(CTD) were made in 20 mM HEPES, pH 7.2, 140 mM KCl and 5% (w/v) glycerol in the presence of 40 nM (Sld5) or 50 nM (Pol1) fluorescein-labelled peptide. Each data point is the mean of three independent experiments. Curve fitting to the experimental data was performed in ProFit 6.2 (QuantumSoft) using a Robust fitting algorithm in combination with a Lorentzian error distribution analysis.

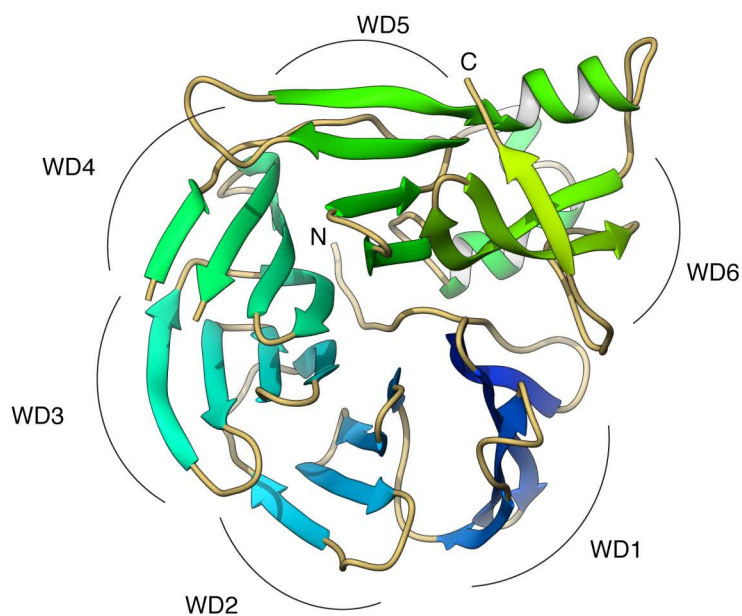
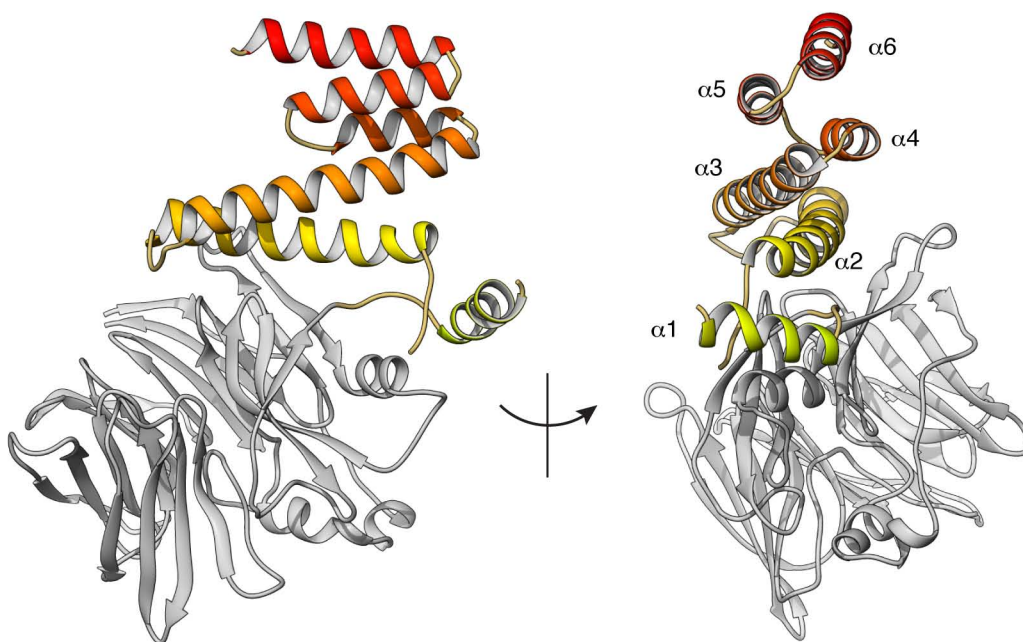
**Native mass spectrometry.** In preparation for non-denaturing nano-electrospray ionization mass spectrometry (native mass spectrometry), protein samples were subjected to two successive rounds of buffer exchange into 500 mM ammonium acetate using illustra NAP-5 columns (GE Healthcare). For reconstitution of the Ctf4(CTD)–Pol  $\alpha$  and Ctf4(CTD)–Sld5 complexes, Ctf4(CTD) was incubated with a tenfold or fivefold molar excess of Pol1 peptide 137-IDNFDDILGEFES-149 or Sld5 peptide 1-MDINIDDILAELDKETTAV-19, respectively, for 30 min before buffer exchange. After buffer exchange, samples were concentrated to at least 50  $\mu\text{M}$  in preparation for mass spectrometric analysis. Native mass spectra were recorded on a Synapt HDMS instrument (Waters), and calibrated using caesium iodide (100 mg ml<sup>-1</sup>) as described previously<sup>44</sup>. Typical parameter values were: capillary voltage 1.8 kV, cone voltage 40–80 V, cone gas 40 l h<sup>-1</sup>, extractor 1.2–2.2 V, ion transfer stage pressure 3.61–3.44 mbar, trap collision energy 10–15 V, transfer collision energy 10–20.0 V, trap and transfer pressure 5.29–5.33  $\times 10^{-2}$  mbar, IMS pressure 5.01–5.02  $\times 10^{-1}$  mbar, TOF analyser pressure 1.17–1.18  $\times 10^{-6}$  mbar. Micromass MassLynx 4.1 was used for data acquisition and processing.

**Artwork.** All structural drawings were prepared with UCSF Chimera<sup>45</sup>.

28. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363–371 (2009).

29. Peränen, J., Rikonen, M., Hyvonen, M. & Kaariainen, L. T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal. Biochem.* **236**, 371–373 (1996).
30. Aslanidis, C. & de Jong, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* **18**, 6069–6074 (1990).
31. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnol.* **17**, 1030–1032 (1999).
32. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124 (1993).
33. Kastner, B. *et al.* GraFix: sample preparation for single-particle electron cryomicroscopy. *Nature Methods* **5**, 53–55 (2008).
34. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
35. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67**, 282–292 (2011).
36. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
37. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
38. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
39. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
40. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
41. Costa, A. *et al.* Structural studies of the archaeal MCM complex in different functional states. *J. Struct. Biol.* **156**, 210–219 (2006).
42. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
43. Kilkenny, M. L., De Piccoli, G., Perera, R. L., Labib, K. & Pellegrini, L. A conserved motif in the C-terminal tail of DNA polymerase  $\alpha$  tethers primase to the eukaryotic replisome. *J. Biol. Chem.* **287**, 23740–23747 (2012).
44. Hernández, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature Protocols* **2**, 715–726 (2007).
45. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

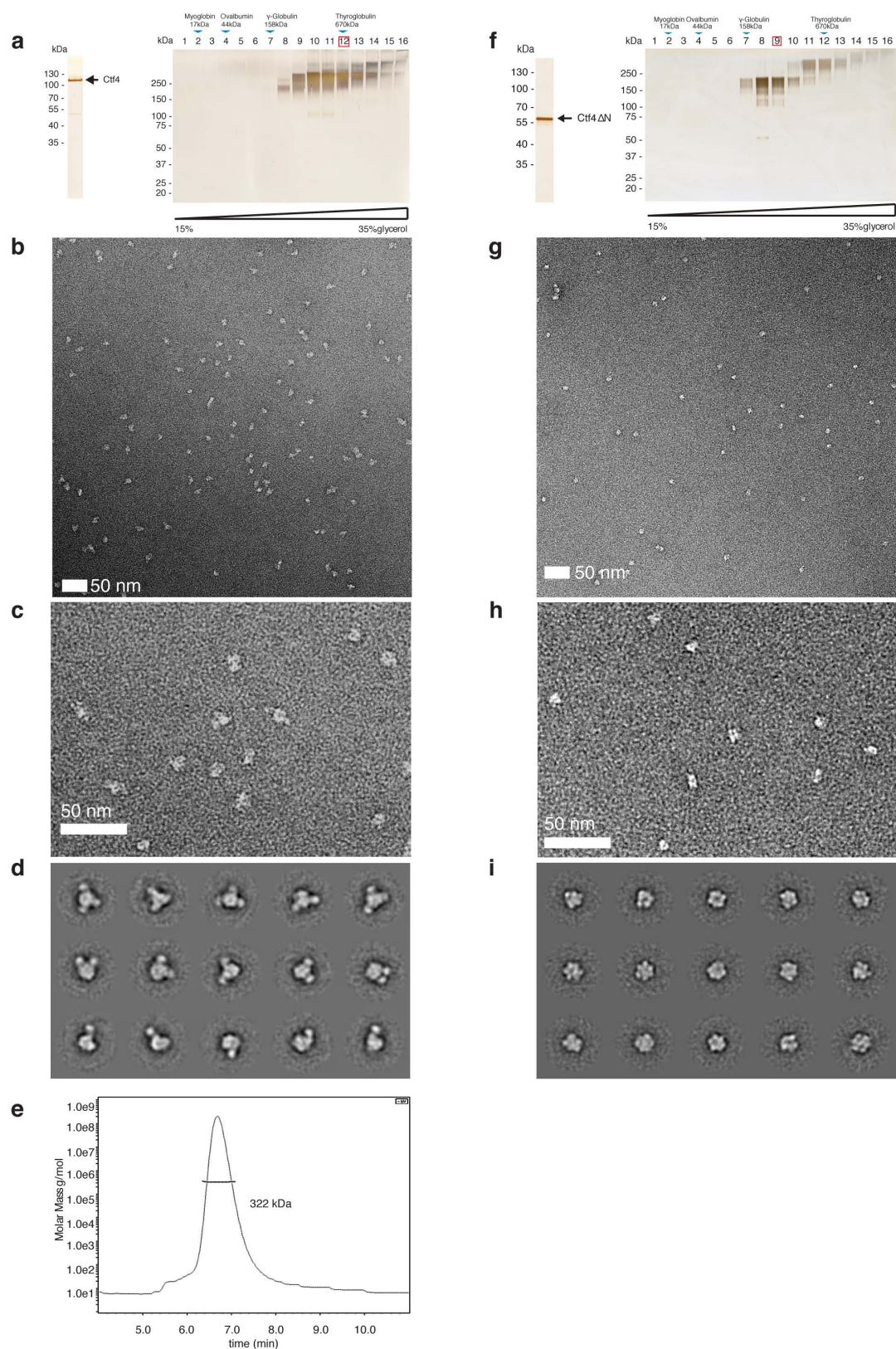


**a****b**

**Extended Data Figure 1 | Crystal structure of yeast Ctf4 region spanning amino acids 471–927 (C-end; Ctf4(CTD)).** **a**, The six-bladed  $\beta$ -propeller of Ctf4(CTD). The structure is drawn as a ribbon diagram, coloured blue to light green from the N to the C terminus. The six blades of the propeller are labelled WD1 to WD6, from the N to C end. The helical domain has been

omitted for clarity. **b**, Two views of the Ctf4(CTD) protomer, highlighting the helical domain of Ctf4(CTD), coloured yellow to red from the amino to the carboxy terminus. The  $\alpha$ -helices are labelled  $\alpha 1$  to  $\alpha 6$ . The  $\beta$ -propeller domain is in light grey.

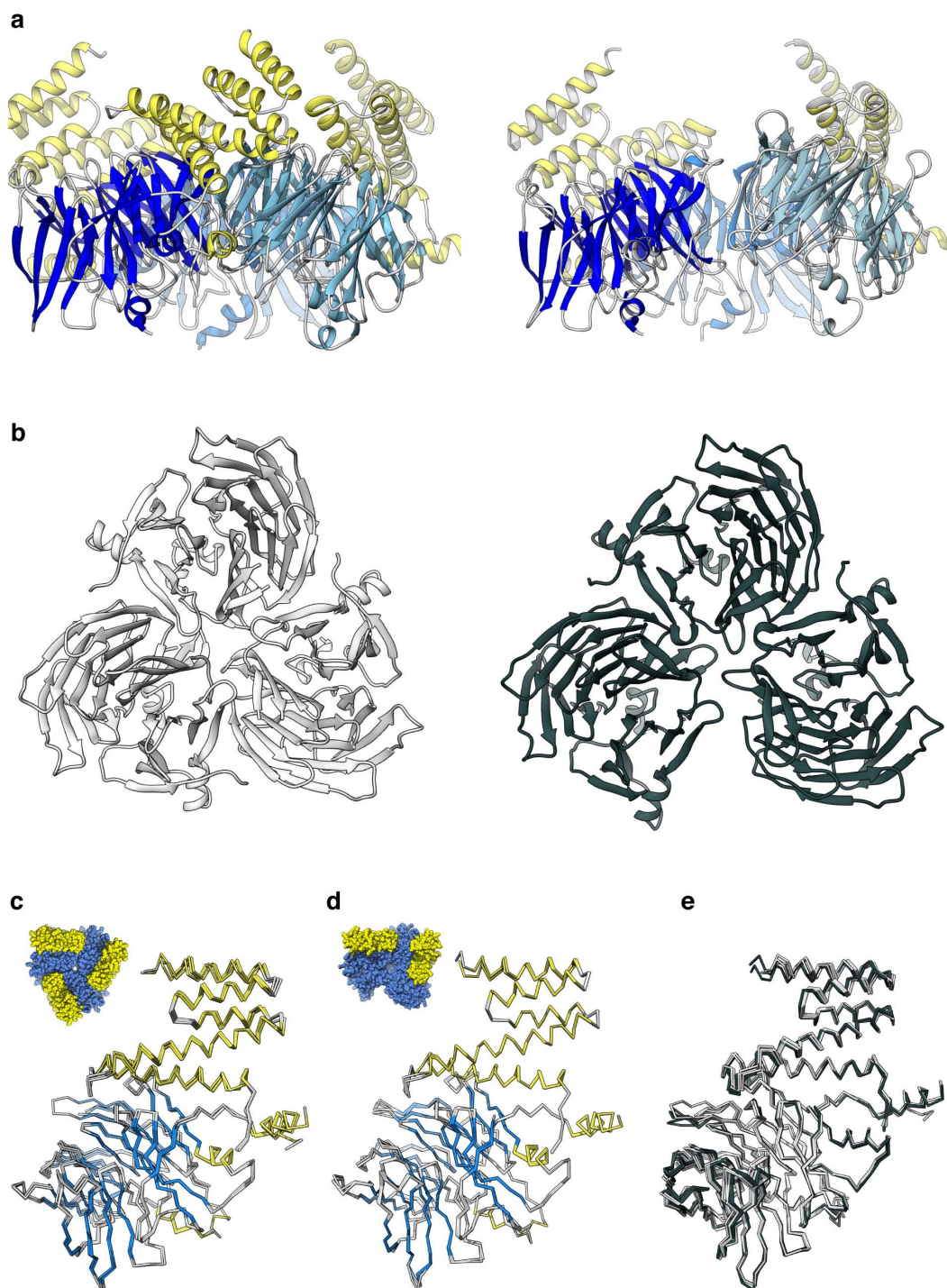




### Extended Data Figure 2 | 2D EM of full-length Ctf4 and Ctf4(CTD).

**a**, Silver-stained SDS-PAGE gel showing the purified full-length yeast Ctf4 and GraFix gel of the same preparation. A red box highlights the fraction imaged by EM. **b**, Representative micrograph of the full-length yeast Ctf4. **c**, Zoomed-in view of the micrograph in panel **b**. **d**, Reference-free class averages of the full-length protein complex highlight the presence of a ring-shaped core linked to mobile, satellite densities. Box size 448 Å. **e**, Multi-angle light scattering reveals that Ctf4 forms stable homotrimers in solution (absolute molecular

mass equates to 322 kDa, with a 0.6% error; expected molecular mass for a monomer: 106.9 kDa). **f**, Silver-stained SDS-PAGE gel showing the purified Ctf4(CTD) and GraFix gel of the same preparation. A red box highlights the fraction imaged by EM. **g**, Representative micrograph of the Ctf4(CTD) complex. **h**, Zoomed-in view of the micrograph in panel **g**. **i**, Reference-free class averages of the Ctf4(CTD) protein complex highlight the presence of an oligomerization core. Box size 448 Å.



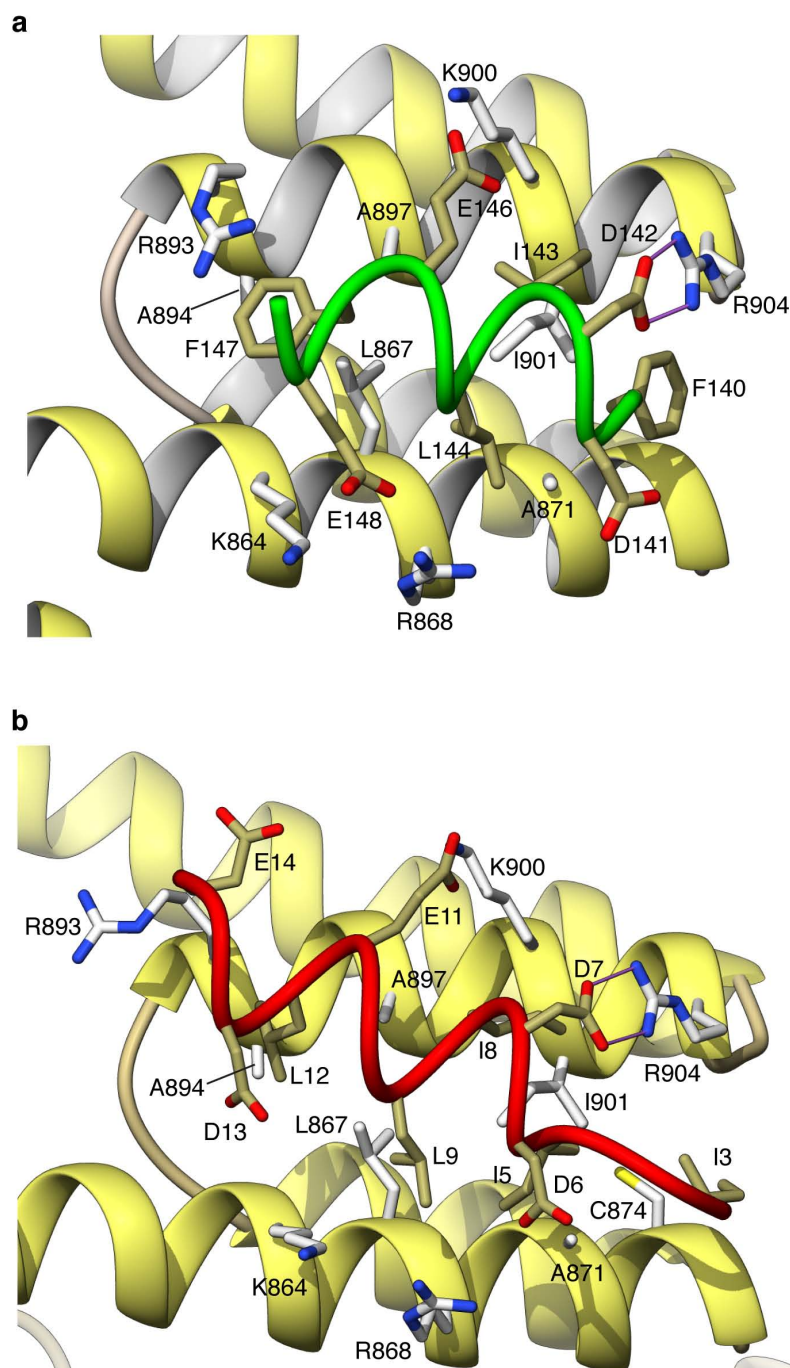
### Extended Data Figure 3 | Crystal forms of the Ctf4(CTD) structure.

**a, b**, Side-by-side comparisons of the symmetric, closed form of the Ctf4(CTD) trimer (left-hand side) with the pseudo-symmetric, open form (right-hand side), used for study of the interaction with the Ctf4-binding motifs of Pol  $\alpha$  and Sld5. Panel **a** shows a side view of the two crystal forms, drawn as ribbons, with the  $\beta$ -propeller domain of the three protomers coloured blue, cyan and light blue, respectively, and the helical domain coloured yellow. Panel **b** shows a top view of the two crystal forms, drawn as ribbons, with the symmetric, closed form of the Ctf4(CTD) trimer in light grey and the pseudo-symmetric, open

form in dark grey. The helical domains have been removed for clarity.

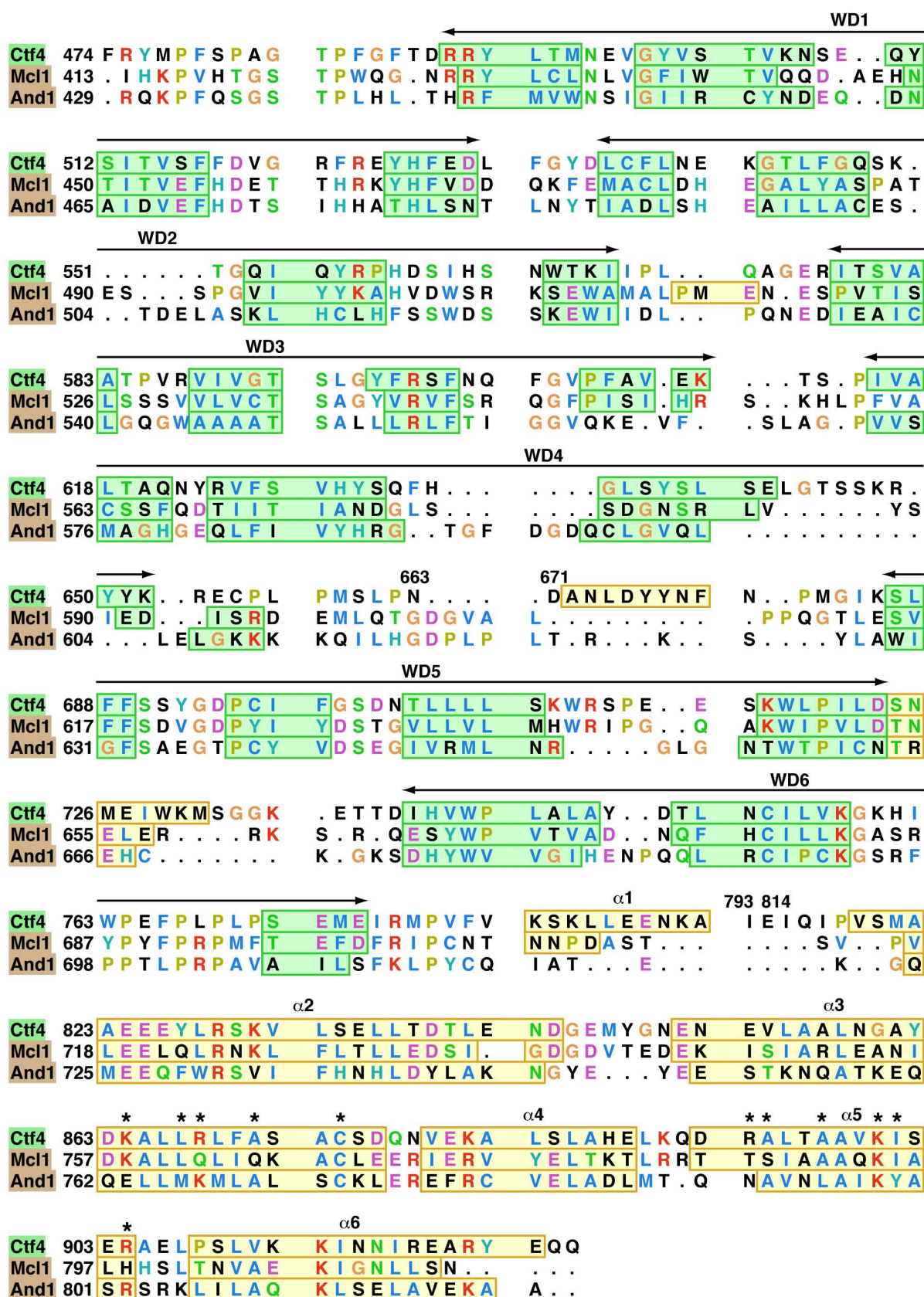
**c–e**, Superpositions of the C $\alpha$  traces of the protomers of the symmetric, closed form of the Ctf4(CTD) trimer, the protomers of the pseudo-symmetric, open Ctf4(CTD) trimer and all protomers of the two crystal forms, respectively. In panels **c** and **d**, the protomers are coloured according to secondary structure, with  $\beta$ -strands in cyan and  $\alpha$ -helices in yellow; the crystal form of the superimposed protomers is highlighted by a space-fill model of the structure in the top-left corner of the panel. In panel **e**, the Ctf4(CTD) protomers of the closed and open forms are coloured light and dark grey, respectively.





**Extended Data Figure 4 | Atomic details of the Ctf4–Pol  $\alpha$  (a) and Ctf4–Sld5 (b) interfaces.** Ctf4(CTD) is drawn as a yellow ribbon, the Ctf4-binding motifs of Pol  $\alpha$  and Sld5 as green and red tubes. The side chains of amino acids at the interface are shown as sticks, with carbon atoms coloured

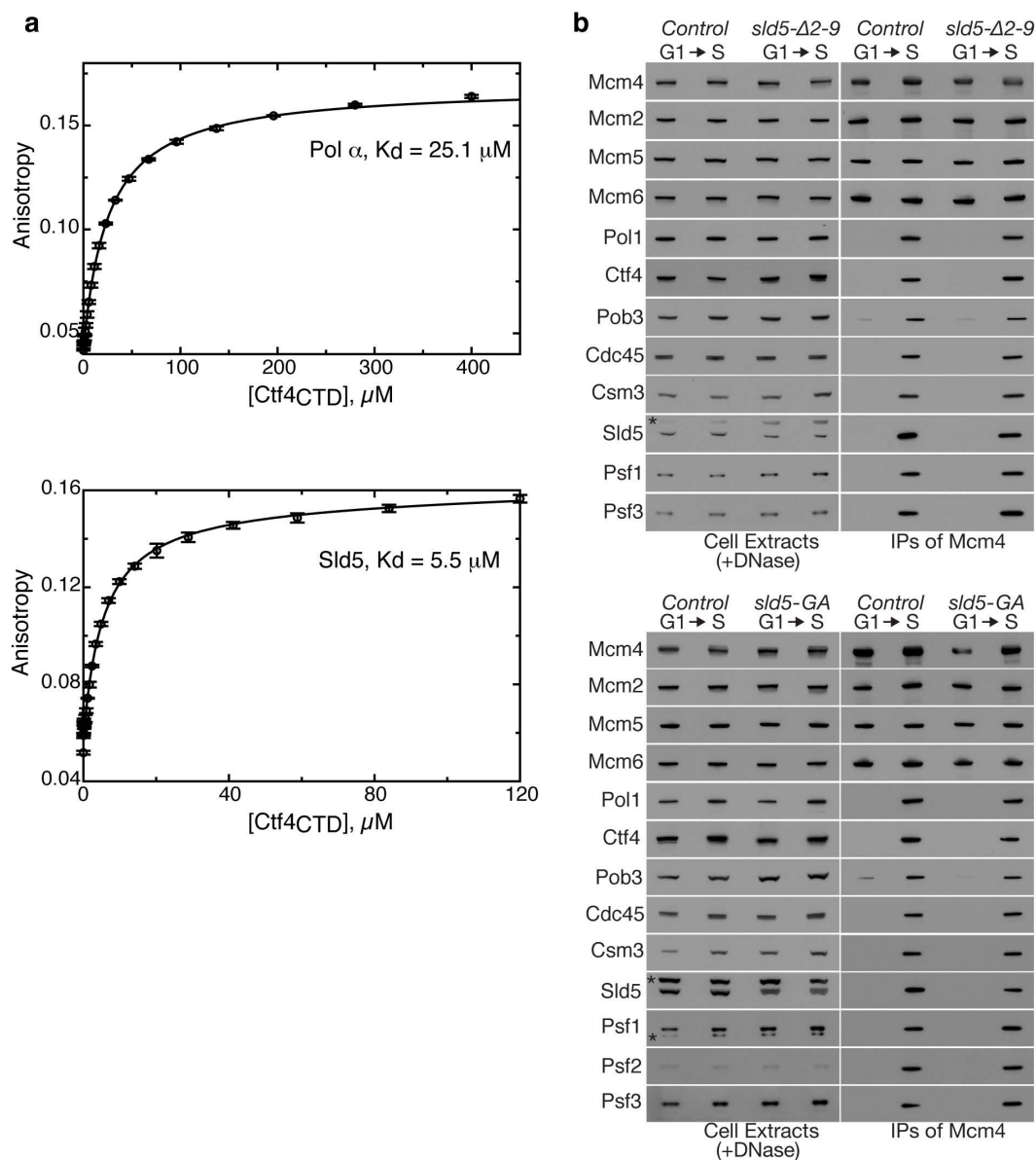
white (Ctf4) or light brown (Pol  $\alpha$  and Sld5), oxygen atoms in red, nitrogen atoms in blue and sulphur atoms in yellow. The bidentate salt link between D142 (Pol1) or D7 (Sld5) and R904 of Ctf4 is shown as solid pink lines.



Extended Data Figure 5 | Structure-based multiple sequence alignment of *S. cerevisiae* Ctf4, *S. pombe* Mcl1 and *H. sapiens* And1. Only the region corresponding to the crystal structure described in the paper is reported. Observed secondary structure elements of Ctf4 and predicted secondary structure elements of Mcl1 and And1 are boxed and shaded in green and yellow

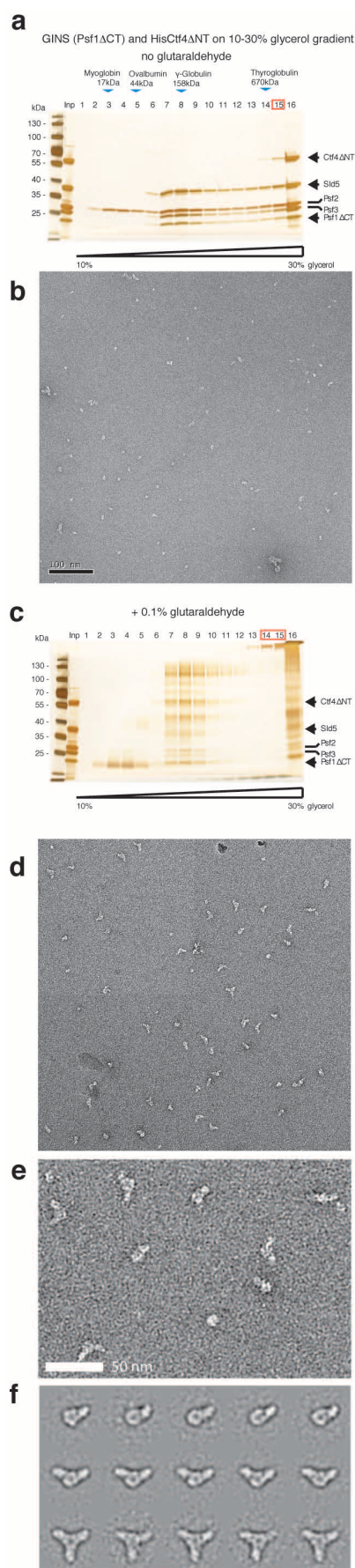
for  $\beta$ -sheets and  $\alpha$ -helices, respectively. The extent of the WD40 domains of the six-bladed  $\beta$ -propeller and the  $\alpha$ -helices of the helical domain are illustrated above the alignment. Ctf4 residues that form the interface with Pol  $\alpha$  and Sld5 are marked by an asterisk.



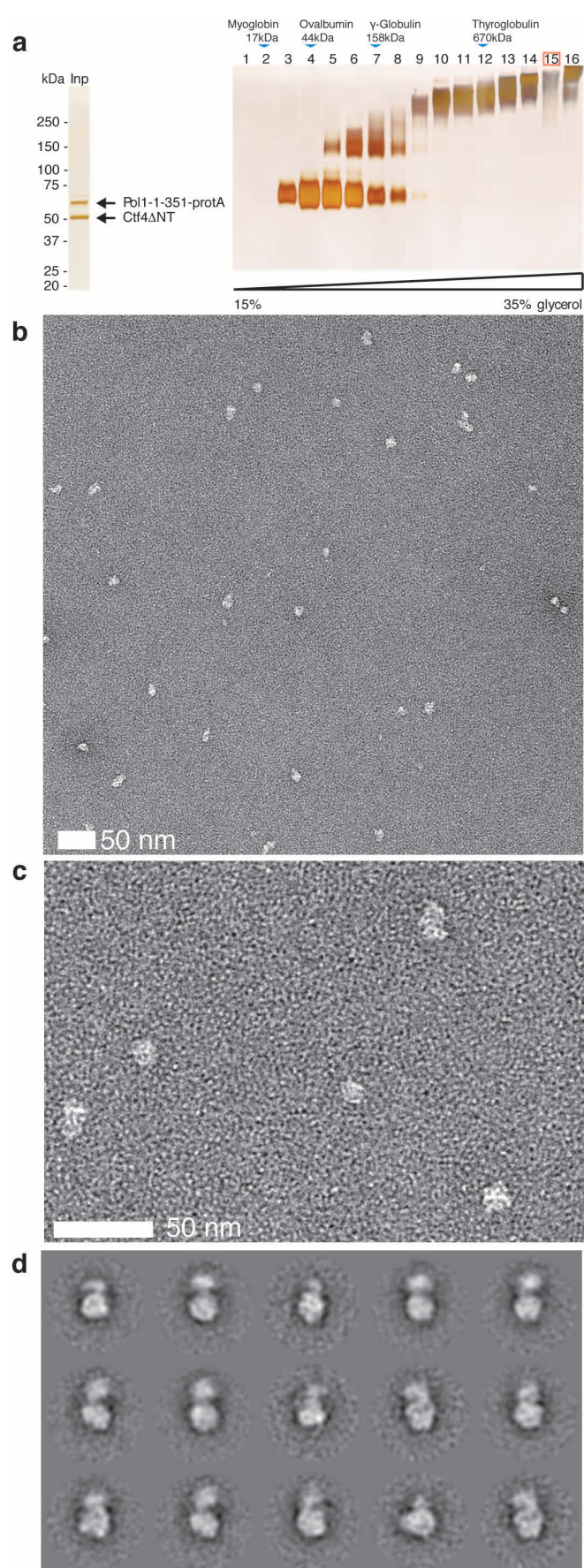


**Extended Data Figure 6 | Ctf4 interactions with Pol  $\alpha$  and CMG.** **a**, Binding affinity of Ctf4(CTD) for the Ctf4-binding motifs of Pol  $\alpha$  (top panel) and Sld5 (bottom panel). Affinity was measured by fluorescence anisotropy of fluorescein-labelled peptides in the presence of increasing amounts of Ctf4. See Methods for experimental details. **b**, CMG still associates with Ctf4 and Pol1 in yeast cells with mutations in the Ctf4-binding motif of Sld5. The budding yeast strains *MCM4-5FLAG* (Control) and *MCM4-5FLAG sld5-Δ2-9* (the endogenous copy of *SLD5* was modified to create *sld5-Δ2-9*, such that the

encoded protein lacks amino acids 2-9) were grown at 24 °C, arrested in G1 phase, and then released into S phase for 30 min. Mcm4-5FLAG was then isolated from cell extracts by immunoprecipitation, and the indicated proteins were detected by immunoblotting with the corresponding antibodies<sup>23</sup> (top panel). An analogous experiment was performed with *MCM4-5FLAG* (Control) and *MCM4-5FLAG sld5-GA* (the endogenous copy of *SLD5* was modified to create *sld5-GA*, such that amino acids 5-9 were changed from Ile-Asp-Asp-Ile-Leu to Gly-Ala-Gly-Ala-Gly) (bottom panel).

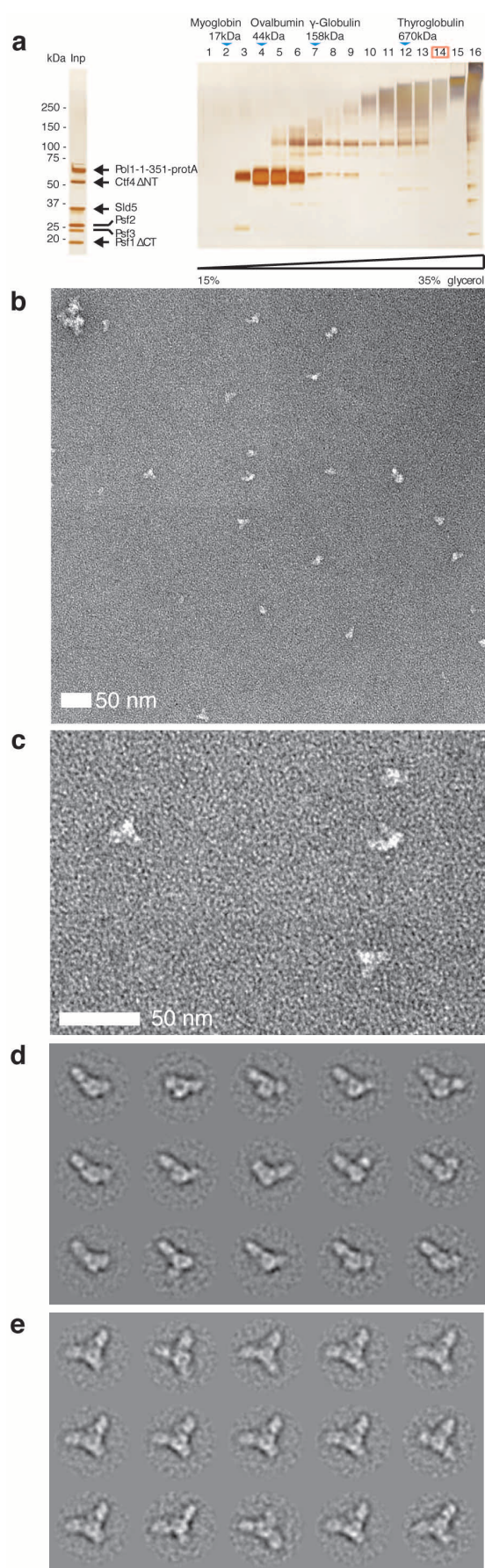


**Extended Data Figure 7 | 2D EM analysis of Ctf4(CTD)–GINS complexes.** Comparison between GINS–Ctf4(CTD) complexes prepared by glycerol gradient or GraFix. **a**, SDS–PAGE gel of non-crosslinked GINS–Ctf4(CTD) complex. A red box highlights the fraction imaged by EM. **b**, Representative micrograph for the non-crosslinked preparation highlights small, globular particles. **c**, SDS–PAGE gel of crosslinked GINS–Ctf4(CTD) complex. A red box highlights the fraction imaged by EM. **d**, Representative micrograph for the crosslinked preparation highlights elongated features compatible with one, two or three GINS docked onto a Ctf4 trimerization core. **e**, Zoomed-in view of the same micrograph. **f**, Representative class averages of the Ctf4(CTD)–GINS complex show a mixture of complexes with clearly discernible stoichiometry: Ctf4<sub>3</sub>–GINS, Ctf4<sub>3</sub>–(GINS)<sub>2</sub>, Ctf4<sub>3</sub>–(GINS)<sub>3</sub>. The box size is 448 Å.



**Extended Data Figure 8 | 2D EM analysis of the Ctf4(CTD)-Pol1(NTD) complex.** **a**, Silver-stained SDS-PAGE gel showing the purified Ctf4(CTD)-Pol1(NTD) complex and GraFix gel of the same preparation. A red box highlights the fraction imaged by EM. **b**, Representative micrograph of the Ctf4(CTD)-Pol1(NTD) complex. **c**, Zoomed-in view of the same micrograph. **d**, Reference-free class averages of the Ctf4(CTD)-Pol1(NTD) complex highlight the presence of a ring-shaped core linked to one mobile globular density. The box size is 448 Å.





**Extended Data Figure 9 | 2D EM analysis of Ctf4(CTD)-GINS-Pol1(NTD) complexes.** **a**, Silver-stained SDS-PAGE gel showing the purified GINS-Ctf4(CTD)-Pol1(NTD) complex and GraFix gel of the same preparation. A red box highlights the fraction imaged by EM. **b**, Representative micrograph of the GINS-Ctf4(CTD)-Pol1(NTD) complex. **c**, Zoomed-in view of the same micrograph. **d**, **e**, Reference-free class averages of the complex highlight the presence of a ring-shaped core linked to 2 or 3 peripheral features. The peripheral features fall into two categories: either elongated radially departing features similar to those seen for the Ctf4(CTD)-GINS complex, or smaller globular densities that were assigned to Pol1(NTD). The box size is 448 Å.



Extended Data Table 1 | X-ray diffraction data processing and crystallographic refinement

	Native	Seleno-methionine	Pol $\alpha$ soak	Sld5 soak
<b>Data collection</b>				
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	89.1 100.3 219.5	107.3, 118.4, 155.7	89.0 100.0 219.5	88.9, 99.7, 218.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
		<i>Peak</i>		
Wavelength (Å)	0.9173	0.97944	0.9795	0.9795
Resolution (Å)	49.17-3.0 (3.16-3.0)*	49.9-2.69 (2.77-2.69)	49.2-2.69 (2.77-2.69)	49.07-2.69 (2.77-2.69)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	0.194 (0.983)	0.228 (3.383)	0.163 (1.473)	0.174 (1.691)
<i>I</i> / $\sigma$ <i>I</i>	12.6 (2.5)	17.4 (1.1)	10.2 (1.4)	10.3 (1.4)
Completeness (%)	99.9 (99.9)	99.5 (94.5)	99.6 (95.0)	99.2 (90.5)
Redundancy	9.3 (9.7)	26.7 (23.9)	7.3 (7.1)	7.3 (7.1)
<b>Refinement</b>				
Resolution (Å)	48.87-3.0	49.19-2.69	49.2-2.694	47.92-2.694
No. reflections	40149	55384	54707	54193
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.1674/0.2049	0.1895/0.2248	0.1718/0.2099	0.1787/0.2141
No. atoms	9447	10421	9617	9684
Protein	9349	10258	9457	9501
Ligand/ion	-	-	-	-
Water	98	163	160	183
B-factors				
Protein	51.6	73.4	59.2	59.5
Ligand/ion	51.8	73.7	59.5	59.8
Water	-	-	-	-
	33.5	56	44.2	47.1
R.m.s deviations				
Bond lengths (Å)	0.003	0.003	0.003	0.002
Bond angles (°)	0.65	0.65	0.67	0.63

\*Statistics for the highest resolution shell is shown in parenthesis.

# Quantitative flux analysis reveals folate-dependent NADPH production

Jing Fan<sup>1\*</sup>, Jiangbin Ye<sup>2\*</sup>, Jurre J. Kamphorst<sup>1</sup>, Tomer Shlomi<sup>1,3</sup>, Craig B. Thompson<sup>2</sup> & Joshua D. Rabinowitz<sup>1</sup>

ATP is the dominant energy source in animals for mechanical and electrical work (for example, muscle contraction or neuronal firing). For chemical work, there is an equally important role for NADPH, which powers redox defence and reductive biosynthesis<sup>1</sup>. The most direct route to produce NADPH from glucose is the oxidative pentose phosphate pathway, with malic enzyme sometimes also important<sup>2,3</sup>. Although the relative contribution of glycolysis and oxidative phosphorylation to ATP production has been extensively analysed, similar analysis of NADPH metabolism has been lacking. Here we demonstrate the ability to directly track, by liquid chromatography–mass spectrometry, the passage of deuterium from labelled substrates into NADPH, and combine this approach with carbon labelling and mathematical modelling to measure NADPH fluxes. In proliferating cells, the largest contributor to cytosolic NADPH is the oxidative pentose phosphate pathway. Surprisingly, a nearly comparable contribution comes from serine-driven one-carbon metabolism, in which oxidation of methylene tetrahydrofolate to 10-formyl-tetrahydrofolate is coupled to reduction of NADP<sup>+</sup> to NADPH. Moreover, tracing of mitochondrial one-carbon metabolism revealed complete oxidation of 10-formyl-tetrahydrofolate to make NADPH. As folate metabolism has not previously been considered an NADPH producer, confirmation of its functional significance was undertaken through knockdown of methylenetetrahydrofolate dehydrogenase (*MTHFD*) genes. Depletion of either the cytosolic or mitochondrial MTHFD isozyme resulted in decreased cellular NADPH/NADP<sup>+</sup> and reduced/oxidized glutathione ratios (GSH/GSSG) and increased cell sensitivity to oxidative stress. Thus, although the importance of folate metabolism for proliferating cells has been long recognized and attributed to its function of producing one-carbon units for nucleic acid synthesis, another crucial function of this pathway is generating reducing power.

Previous examination of NADPH production during cell growth has analysed metabolic fluxes in cells using <sup>13</sup>C and <sup>14</sup>C isotope tracers<sup>4–7</sup>. For NADPH metabolism, however, carbon tracers alone are insufficient, because they cannot determine whether a particular redox reaction is making NADH versus NADPH or the reaction's fractional contribution to total cellular NADPH production. To address these limitations, we developed a deuterium tracer approach that directly measures NADPH redox active hydrogen labelling. To probe the oxidative pentose phosphate pathway, we shifted cells from unlabelled to [1-<sup>2</sup>H]glucose or [3-<sup>2</sup>H]glucose (Fig. 1a) and measured the resulting NADP<sup>+</sup> and NADPH labelling by liquid chromatography–mass spectrometry (LC–MS)<sup>8</sup>, as shown in the mass spectrum in Fig. 1b (for associated chromatogram, see Extended Data Fig. 1a). The M+1 and M+2 peaks in NADP<sup>+</sup> are natural isotope abundance, primarily from <sup>13</sup>C. The difference between NADP<sup>+</sup> and NADPH reflects the redox active hydrogen labelling. The labelling of NADPH's redox-active hydrogen is fast (*t*<sub>1/2</sub> ~ 5 min) (Fig. 1c; note, as opposed to relative mass intensities, all fractional labelling data are corrected for natural isotope abundance). NADPH labelling was similar across four different transformed mammalian cell lines. Knockdown of the committed enzyme of the oxidative pentose phosphate pathway,

glucose-6-phosphate dehydrogenase, eliminated most of the labelling, confirming that the NADPH-deuterium labelling reflects oxidative pentose phosphate pathway flux (Fig. 1d).

Because most NADPH is cytosolic<sup>9</sup>, the <sup>2</sup>H-glucose labelling results can be used to quantitate the fractional contribution of the oxidative pentose phosphate pathway (oxPPP) to total cytosolic NADPH production:

$$\text{Fraction}_{\text{NADPH from oxPPP}} = 2 \times \left( \frac{[^2\text{H}]\text{NADPH}}{\text{Total NADPH}} \right) \times \left( \frac{[^2\text{H}]\text{G6P}}{\text{Total G6P}} \right)^{-1} \times C_{\text{KIE}} \quad (1)$$

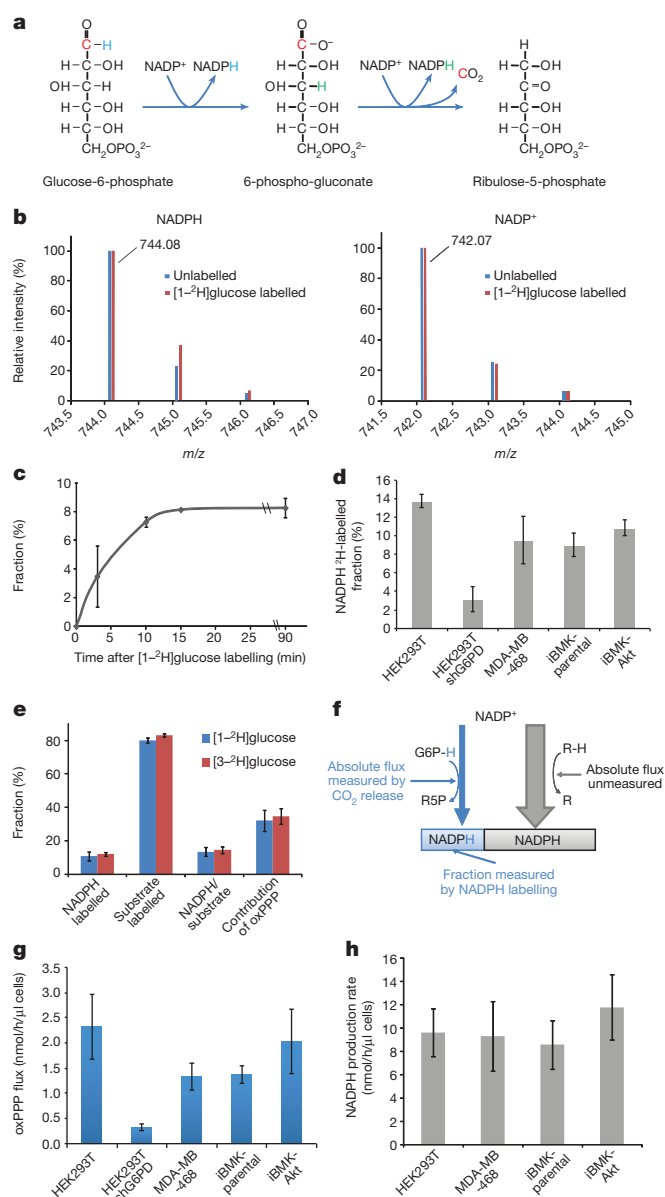
The terms in parentheses are the fractional <sup>2</sup>H-labelling of NADPH's redox active hydrogen and of glucose-6-phosphate's targeted hydrogen (Fig. 1e, Extended Data Fig. 1b–d). The term *C*<sub>KIE</sub> accounts for the deuterium kinetic isotope effect<sup>10,11</sup> (see Methods, Extended Data Fig. 1e–g). Note that these <sup>2</sup>H-labelling experiments directly measure the fraction of NADPH made by the oxidative pentose phosphate pathway without relying on measurement of the absolute pathway flux. Using either [1-<sup>2</sup>H]glucose or [3-<sup>2</sup>H]glucose, we find that oxidative pentose phosphate pathway accounts for 30–50% of overall NADP<sup>+</sup> reduction.

The inferred fractional contribution of oxidative pentose phosphate pathway to NADPH production can be used to deduce the total cytosolic NADPH production rate, which is equal to the absolute oxidative pentose phosphate pathway flux divided by the fractional contribution of the oxidative pentose phosphate pathway to NADPH production (Fig. 1f). To this end, we measured absolute oxidative pentose phosphate pathway flux using two orthogonal approaches. The first approach measures <sup>14</sup>CO<sub>2</sub> release from [1-<sup>14</sup>C]glucose versus [6-<sup>14</sup>C]glucose (Extended Data Figs 2a–c and 3). The second measures the kinetics of 6-phosphogluconate labelling from uniformly <sup>13</sup>C-labeled glucose ([U-<sup>13</sup>C]glucose) (Extended Data Fig. 2d–f). Both approaches gave consistent fluxes, with the radioactive measurement being more precise (Extended Data Fig. 2g). As confirmation of its specificity, we knocked down glucose-6-phosphate dehydrogenase and observed markedly reduced oxidative pentose phosphate pathway <sup>14</sup>CO<sub>2</sub> release (Fig. 1g). In the absence of such knockdown, the observed oxidative pentose phosphate pathway flux ranged from 1–2.5 nmol μl<sup>−1</sup> h<sup>−1</sup> (in which volume is the packed cell volume; Fig. 1g). This flux is similar to, but slightly less than, the cellular ribose demand (Extended Data Fig. 3f). In combination with the fractional NADPH labelling, we deduced a total cytosolic NADPH production rate of ~10 nmol μl<sup>−1</sup> h<sup>−1</sup> (Fig. 1h), which is 5–20% of the glucose uptake rate.

To investigate whether we could use <sup>2</sup>H-labelling to directly observe NADPH production by other pathways (Fig. 2a), we administered [2,3,4,4-<sup>2</sup>H]glutamine and [2,3,3-<sup>2</sup>H]aspartate to cells. Downstream products of glutamine can potentially transfer <sup>2</sup>H to NADPH via glutamate dehydrogenase or malic enzyme, whereas downstream products of aspartate may do so via isocitrate dehydrogenase (Extended Data Fig. 4a–f). We observed identical mass spectra for NADP<sup>+</sup> and NADPH after feeding the deuterium-labelled glutamine and aspartate (Fig. 2b, c and Extended Data Fig. 4b, d), and thus could not directly assign a fractional

<sup>1</sup>Department of Chemistry and Lewis Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08540, USA. <sup>2</sup>Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>3</sup>Department of Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel.

\*These authors contributed equally to this work.



**Figure 1 | Quantification of NADPH labelling via the oxidative pentose phosphate pathway and of total cytosolic NADPH production.** **a**, Oxidative pentose phosphate pathway schematic. **b**, Mass spectra of NADPH and NADP<sup>+</sup> from cells labelled with [1-<sup>2</sup>H]glucose (iBMK-parental cells, 20 min). **c**, Kinetics of NADPH labelling from [1-<sup>2</sup>H]glucose (iBMK-parental cells). **d**, NADPH labelling from [1-<sup>2</sup>H]glucose (20 min). **e**, [1-<sup>2</sup>H]glucose and [3-<sup>2</sup>H]glucose yield similar NADPH labelling (iBMK-parental cells, 20 min). Substrate labelling is reported for glucose-6-phosphate for [1-<sup>2</sup>H]glucose and 6-phosphogluconate for [3-<sup>2</sup>H]glucose. oxPPP, oxidative pentose phosphate pathway. **f**, Schematic illustrating that the total cytosolic NADP<sup>+</sup> reduction flux is the absolute oxidative pentose phosphate pathway flux (measured based on <sup>14</sup>CO<sub>2</sub> excretion) divided by the fractional oxidative pentose phosphate pathway contribution (measured based on NADPH <sup>2</sup>H-labelling). **g**, Oxidative pentose phosphate pathway flux based on difference in <sup>14</sup>CO<sub>2</sub> release from [1-<sup>14</sup>C]glucose and [6-<sup>14</sup>C]glucose. **h**, Total cytosolic NADP<sup>+</sup> reduction flux. All results are mean ± s.d., *n* ≥ 2 biological replicates from a single experiment and results were confirmed in multiple experiments.

contribution to these pathways. Given recent evidence that malic enzyme is particularly important in cancer<sup>2,3</sup>, we used an orthogonal approach based on feeding [U-<sup>13</sup>C]glutamine and measuring labelling of pyruvate, and lactate to evaluate its activity (Extended Data Fig. 4g, h). Although such carbon tracer studies cannot distinguish between NADH-dependent and NADPH-dependent malic enzyme, they put an upper bound on their

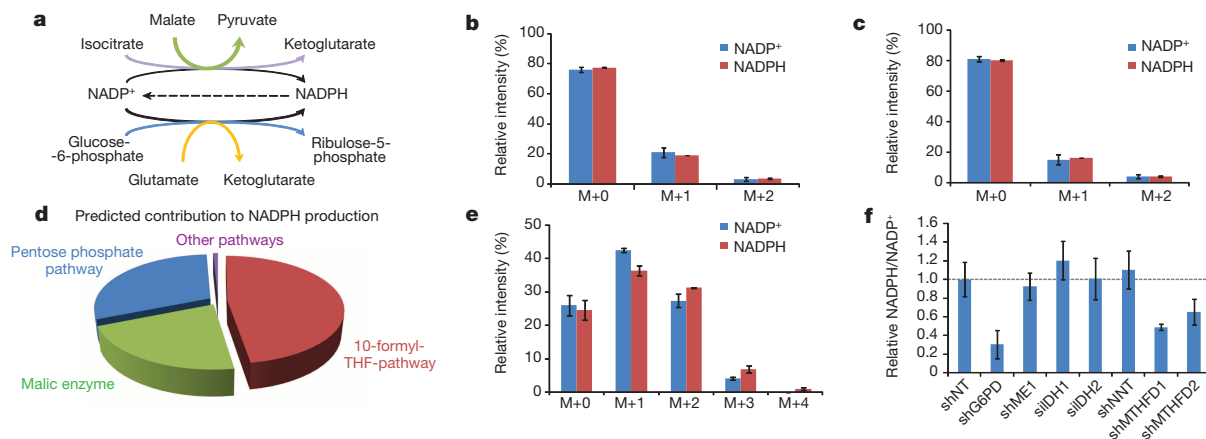
collective activities, which ranged from 15% to 50% of cytosolic NADPH production depending on the cell line.

To identify other potential NADPH producing pathways, we used a genome-scale human metabolic model<sup>12</sup>. We constrained the model based on the observed steady-state growth rate, biomass composition, and metabolite uptake and excretion rates of immortalized baby mouse kidney cells (iBMK-parental cells)<sup>13</sup>, without enforcing any constraints on NADPH production routes. The model, assessed via flux balance analysis with an objective of minimizing total enzyme expression requirements and hence flux<sup>14</sup> (see Methods), predicted that both the oxidative pentose phosphate pathway and malic enzyme contribute ~30% of NADPH production (Fig. 2d). Surprisingly, however, ~40% of NADPH production was predicted to come from one-carbon metabolism mediated by tetrahydrofolate (THF). An alternative objective function of maximizing growth rate further predicts a potentially substantial contribution of folate metabolism to NADPH production (Extended Data Fig. 5a, b).

The main folate-dependent NADPH-producing pathway was predicted to involve transfer of a one-carbon unit from serine to THF, followed by oxidation of the resulting product (methylene-THF) by the enzyme MTHFD to form the purine precursor formyl-THF with concomitant NADPH production. To assess whether this pathway indeed contributes to NADPH production, we supplied cells with [2,3,3-<sup>2</sup>H]serine and observed labelling of both NADP<sup>+</sup> and NADPH. The NADP<sup>+</sup> labelling results from incorporation of the serine-derived formyl-THF one-carbon unit into the adenine ring of NADP<sup>+</sup>. Relative to NADP<sup>+</sup>, the labelling pattern of NADPH was shifted towards more heavily labelled forms, indicating specific labelling of the redox active hydrogen of NADPH (Fig. 2e and Extended Data Fig. 5c, d). Thus, we were able to directly confirm that serine-driven folate metabolism contributes to NADP<sup>+</sup> reduction.

To assess the functional significance of different pathways to NADPH homeostasis, in HEK293T cells we knocked down a variety of potential NADPH-producing enzymes and measured the cellular NADPH/NADP<sup>+</sup> ratio (Fig. 2f). Although knockdown of malic enzyme 1 (*ME1*), cytosolic or mitochondrial NADP-dependent isocitrate dehydrogenase (*IDH1* and *IDH2*), and transhydrogenase (*NNT*) did not significantly impact NADPH/NADP<sup>+</sup>, knockdown of glucose-6-phosphate dehydrogenase or either isozyme of methylene tetrahydrofolate dehydrogenase (*MTHFD1*, cytosolic, or *MTHFD2*, mitochondrial) substantially decreased it. These observations further support the primacy, at least in this growing cell line, of the pentose phosphate and folate pathways in NADPH production.

The importance of both isozymes of methylene tetrahydrofolate dehydrogenase suggests that cytosolic and mitochondrial folate metabolism (Fig. 3a) both contribute to NADPH homeostasis. The product of methylene tetrahydrofolate dehydrogenase, 10-formyl-THF, is a required purine precursor, with each purine ring containing two formyl groups. Thus, the cytosolic 10-formyl-THF production rate must be at least twice the purine biosynthetic flux. The most direct path to cytosolic 10-formyl-THF is via MTHFD1 with concomitant NADPH production (Fig. 3a, solid blue lines). Alternatively, 10-formyl-THF could potentially be made from formate initially generated in the mitochondrion (Fig. 3a, dashed lines)<sup>15,16</sup>. To distinguish between these possibilities, we administered [U-<sup>13</sup>C]glycine, which contributes selectively to mitochondrial one-carbon pools (Fig. 3a, green lines). Glycine is assimilated intact into purines, resulting in M+2 labelling of ATP; however, we did not observe any M+1, M+3 or M+4 ATP, indicating that mitochondrial-derived one-carbon units do not contribute to purine biosynthesis (Fig. 3b). Consistent with this, supplying [U-<sup>13</sup>C]serine revealed that most one-carbon units assimilated into purines come from serine (Extended Data Fig. 6a, b), and knockdown of *MTHFD1* nearly eliminated NADPH redox-active hydrogen labelling from [2,3,3-<sup>2</sup>H]serine (Fig. 3c). Assuming that all 10-formyl-THF production for purine synthesis is coupled via MTHFD1 to NADP<sup>+</sup> reduction, the total NADPH production rate is ~2 nmol μl<sup>-1</sup> h<sup>-1</sup> (Fig. 3d) or ~20% of total cytosolic NADPH flux. To probe potential further oxidation of serine, we administered [3-<sup>14</sup>C]serine and observed <sup>14</sup>CO<sub>2</sub> release, a result implying that the THF pathway



**Figure 2 | Pathways contributing to NADPH production.** **a**, Canonical NADPH production pathways. **b**, NADPH and  $\text{NADP}^+$  isotopic distribution (without correction for natural isotope abundances) after incubation with [2,3,3,4,4- $^2\text{H}$ ]glutamine tracer to probe NADPH production via glutamate dehydrogenase and malic enzyme (HEK293T cells, 48 h). See also Extended Data Fig. 4. **c**, NADPH and  $\text{NADP}^+$  isotopic distribution as in **b** using [2,3,3- $^2\text{H}$ ]aspartate tracer to probe NADPH production via IDH. See also Extended Data Fig. 4. **d**, NADPH production routes predicted by experimentally constrained genome-scale flux balance analysis. **e**, NADPH and  $\text{NADP}^+$  isotopic distribution as in **b** using the [2,3,3- $^2\text{H}$ ]serine tracer to

probe NADPH production via folate metabolism (no glycine in the media). See also Extended Data Fig. 5. **f**, Relative NADPH/ $\text{NADP}^+$  ratio in HEK293T cells with knockdown of various potential NADPH producing enzymes: glucose-6-phosphate dehydrogenase (G6PD), cytosolic malic enzyme (ME1), cytosolic and mitochondrial isocitrate dehydrogenase (IDH1 and IDH2), transhydrogenase (NNT), and cytosolic and mitochondrial methylene tetrahydrofolate dehydrogenase (MTHFD1 and MTHFD2). Plotted ratios are relative to vector control knockdown. Results are mean  $\pm$  s.d.,  $n \geq 2$  biological replicates from a single experiment and results were confirmed in multiple experiments.

runs in excess of one-carbon demand yielding additional NADPH (Fig. 3d and Extended Data Fig. 7).

We also investigated the consequences of elimination of serine from the medium (Extended Data Fig. 8). As has been observed previously both *in vitro*<sup>17,18</sup> and in tumour models<sup>19</sup>, serine depletion impaired cell growth (Extended Data Fig. 8b). Consistent with one important downstream product of serine being NADPH, its removal decreased NADPH/ $\text{NADP}^+$  (Extended Data Fig. 8c). Glycine is both a product of serine metabolism, and itself a potential source of one-carbon units via the mitochondrial glycine cleavage system, whose expression has been linked to oncogenic transformation<sup>20</sup>. We therefore tested the impact of both removing serine and increasing glycine in the culture media. We found that increased glycine further impaired cell growth and decreased the NADPH/ $\text{NADP}^+$  ratio (Extended Data Fig. 8b, c). These results are consistent with increased glycine impairing methylene-THF production, perhaps due to reverse flux through serine hydroxymethyltransferase (Extended Data Fig. 8d, e).

The above results establish a major contribution of serine-driven one-carbon metabolism in NADPH homeostasis. Knockdown of *MTHFD2* also alters NADPH/ $\text{NADP}^+$ , suggesting an additional role for mitochondrial one-carbon metabolism. Mitochondrial folate-dependent enzymes, especially *MTHFD2*, are overexpressed across human cancers<sup>21</sup>. To probe specifically mitochondrial folate metabolism, we administered  $^{14}\text{C}$ -labelled glycine and monitored radioactive  $\text{CO}_2$  release. The glycine cleavage system releases glycine C1 as  $\text{CO}_2$ , while transferring glycine C2 to THF, making methylene-THF. Notably, almost as much radioactive  $\text{CO}_2$  was released from [2- $^{14}\text{C}$ ]glycine as from [1- $^{14}\text{C}$ ]glycine (Fig. 3e), indicating that a majority of mitochondrial methylene-THF is fully oxidized to  $\text{CO}_2$ . Consistent with such complete oxidation, when we administered  $^{13}\text{C}$ -labelled glycine, we did not observe transfer of one-carbon units to the cytosol based on the thymidine triphosphate (dTTP) or methionine labelling, with dTTP's one-carbon unit coming from serine (90–100%) and methionine coming from the medium (Extended Data Fig. 6c–f). As expected on the basis of the mitochondrial methylene-THF oxidation pathway, release of glycine C2 as  $\text{CO}_2$  was decreased by knockdown of either *MTHFD2* or *ALDH1L2* (Extended Data Fig. 7g). Such complete one-carbon unit oxidation may be beneficial for reducing the cellular glycine concentration. In addition, it produces mitochondrial NADPH. Thus, two functions of mitochondrial folate metabolism are glycine detoxification and NADPH production.

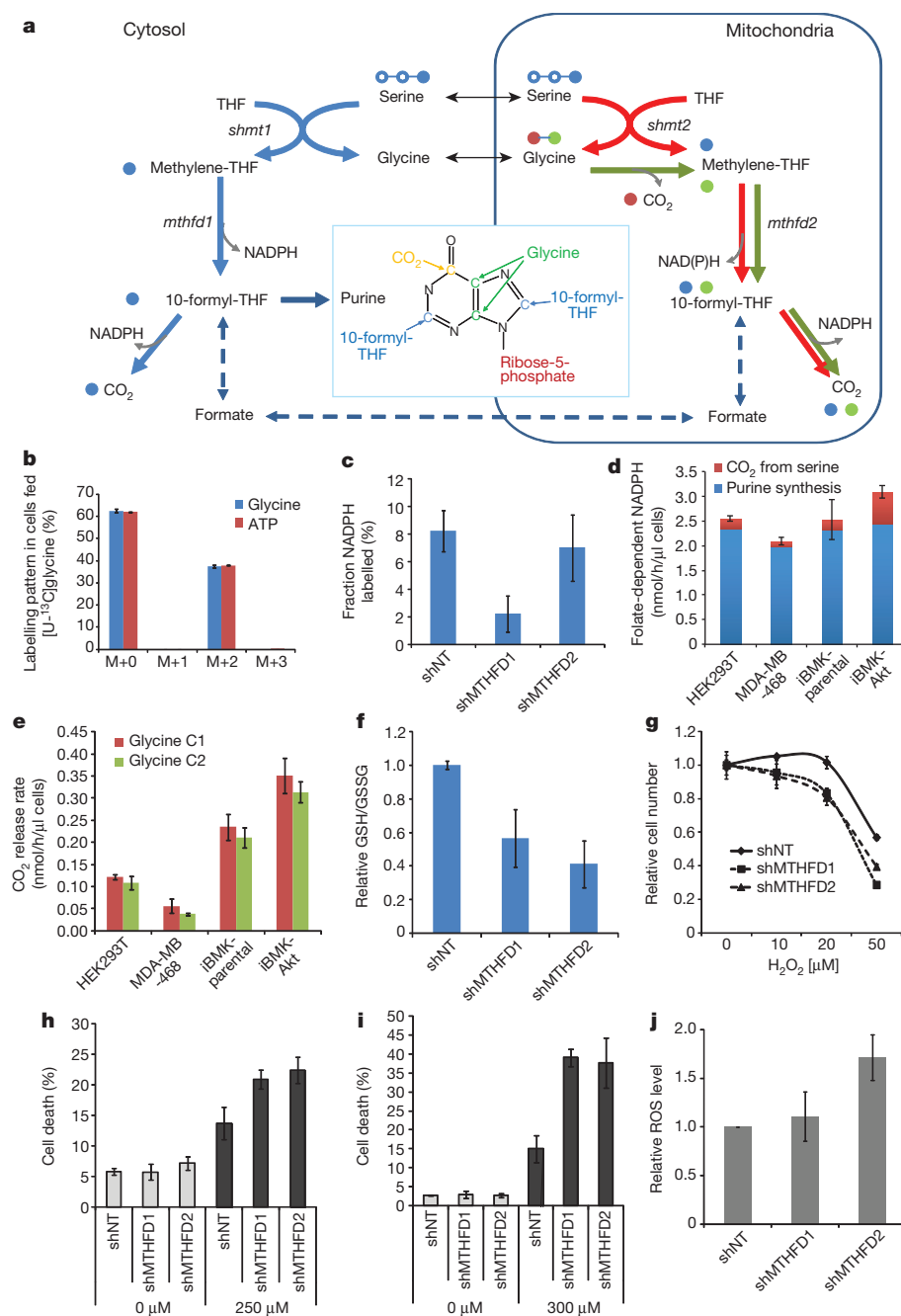
One important role of NADPH is antioxidant defence. Consistent with folate metabolism being a substantial NADPH producer, antifolates have been found to induce oxidative stress<sup>22</sup>. To more directly link folate-mediated NADPH production with cellular redox defenses, we measured glutathione, reactive oxygen species and hydrogen peroxide sensitivity of *MTHFD1* and *MTHFD2* knockdown cells. Knockdown of either isozyme decreased the ratio of reduced to oxidized glutathione (Fig. 3f) and impaired resistance to oxidative stress induced by hydrogen peroxide (Fig. 3g, h) or diamide (Fig. 3i). *MTHFD2* knockdown specifically increased reactive oxygen species (Fig. 3j), and *ALDH1L2* knockdown decreased the ratio of reduced to oxidized glutathione (Extended Data Fig. 7h), demonstrating that the complete mitochondrial methylene-THF oxidation pathway is required for redox homeostasis.

A major open question regards the relative use of NADPH for biosynthesis versus redox defence. To address this, we compared total cytosolic NADPH production (as measured above) to consumption for biosynthesis (Fig. 4a, Methods) based on the measured cellular content of DNA, amino acids and lipids; their production routes (measured by  $^{13}\text{C}$  tracer experiment, see Methods); and cellular growth rate (Extended Data Fig. 9a–g). The overall demand for NADPH for biosynthesis is  $> 80\%$  of total cytosolic NADPH production (Fig. 4b), with most of this NADPH consumed by fatty acid synthesis. At least in transformed cells growing under aerobic conditions, most cytosolic NADPH is devoted to biosynthesis, not redox defence.

To evaluate NADPH consumption for redox defence under overt redox stress, we treated HEK293T cells with hydrogen peroxide at a concentration that blocks growth without causing substantial cell death and measured the total cytosolic NADPH production rate. The rate was  $5.5 \text{ nmol } \mu\text{l}^{-1} \text{ h}^{-1}$ , about half that in freely growing cells (Extended Data Fig. 9h). Thus, consistent with the majority of cytosolic NADPH in growing cells being used for biosynthesis, growth-inhibiting oxidative stress decreases cytosolic NADPH production.

The production of NADPH by the oxidative pentose phosphate pathway, which makes the nucleotide building block ribose, and by the 10-formyl-THF pathway, which contributes to purine synthesis, leads to an inherent coupling of nucleotide synthesis with NADPH production. These reactions together produce in growing cells roughly the amount of NADPH required for replication of cellular lipids (Fig. 4b). Interruption of this intrinsic coordination by feeding of purines can impair cell growth<sup>23</sup>. In non-growing cells, or other cases in which NADPH





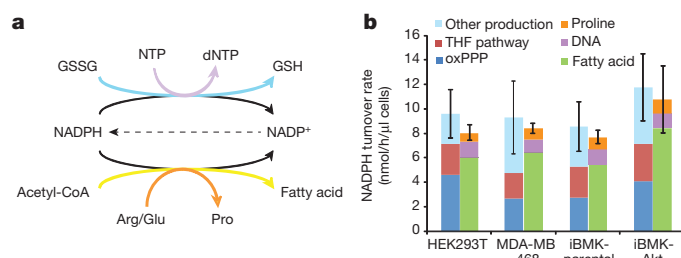
**Figure 3 | Quantification of folate-dependent NADPH production.** **a**, Pathway schematic with serine C3 in blue, glycine C1 in red and glycine C2 in green. **b**, Glycine and ATP labelling pattern after incubation with  $[U-^{13}C]$ glycine (HEK293T cells, 24 h). The lack of M+3 and M+4 ATP indicates that no glycine-derived one-carbon units contribute to purine synthesis. **c**, Fraction of NADPH labelled at the redox active hydrogen after 24 h incubation with  $[2,3,3-^2H]$ serine in HEK293T cells with stable *MTHFD1* or *MTHFD2* knockdown. Same cell lines used also in **f–j**. **d**, Absolute rate of cytosolic folate-dependent NADPH production. **e**,  $CO_2$  release rate from glycine C1 and glycine C2. **f**, GSH/GSSG ratio. **g**, Relative growth, normalized to untreated samples, during 48 h exposure to  $H_2O_2$ . **h**, Cell death after 24 h exposure to 250 μM  $H_2O_2$ . **i**, Cell death after 24 h exposure to 300 μM diamide. **j**, Relative reactive oxygen species (ROS) levels measured using dichlorodihydrofluorescein diacetate (DCFH-DA) assay. Mean  $\pm$  s.d.,  $n = 3$ .

needs outstrip production coupled to nucleotide synthesis, it is likely that alternative pathways, for example, malic enzyme and IDH, will be of greater importance than observed here.

The contribution of the 10-formyl-THF pathway to NADPH production is particularly interesting in light of the importance of metabolism of serine and glycine, the major carbon sources of this pathway, to cancer growth<sup>24</sup>. Serine synthesis is promoted by the cancer-associated M2 isozyme of pyruvate kinase (PKM2) and by amplification of 3-phosphoglycerate dehydrogenase<sup>17,18</sup>. The present data suggest that serine serves dual roles in providing both one-carbon units and NADPH. In this respect, it is intriguing that PKM2, in addition to sensing serine<sup>25,26</sup>, is inactivated by oxidative stress<sup>27</sup>. Such inactivation should increase 3-phosphoglycerate and thus potentially serine-driven NADPH production.

In addition to synthesizing serine, rapidly growing cells avidly consume glycine<sup>28</sup>. Intriguingly, although only intact glycine (and not glycine-derived one-carbon units) is incorporated into purines, knockdown of the glycine cleavage system impairs cancer growth<sup>20</sup>. We find that a

majority of glycine-derived one-carbon units are fully oxidized, arguing against the glycine cleavage system's primary role, at least in the tested cell lines, being to release one-carbon units to the cytosol. Instead,



**Figure 4 | Comparison of NADPH production and consumption.** **a**, Main NADPH consumption pathways. **b**, NADPH production and consumption fluxes. Mean  $\pm$  s.d., with error bar showing the variation of total production or consumption,  $n = 3$ .

its function may be simultaneous elimination of unwanted glycine and production of mitochondrial NADPH.

Understanding NADPH's production and consumption routes is essential to a global understanding of metabolism. The approaches provided here will enable evaluation of these routes in different cell types and environmental conditions. Analogous measurements for ATP, achieved first more than a half century ago<sup>29</sup>, have formed the foundation for much of subsequent metabolism research. Given NADPH's comparable significance in medically important processes including lipogenesis, oxidative stress, and tumour growth<sup>30</sup>, quantitative analysis of its metabolism may prove of similar importance.

## METHODS SUMMARY

Cells were grown in Dulbecco's modified eagle media (DMEM) without pyruvate (CELLGRO) with 10% dialysed fetal bovine serum (Invitrogen) in 5% CO<sub>2</sub> at 37 °C and harvested at ~80% confluency. Stable knockdown cell lines were generated by shRNA-expressing lentivirus with puromycin selection. *IDH1*, *IDH2* and *ALDH1L2* knockdown was generated by transfecting cells with siRNA. For confirmation of knockdown, see Extended Data Fig. 10. For metabolite measurements, metabolism was quenched and metabolites extracted by aspirating media and immediately adding –80 °C 80:20 methanol:water. Supernatants from two rounds of extraction were combined, dried under N<sub>2</sub>, resuspended in water, placed in a 4 °C autosampler, and analysed within 6 h by reversed-phase ion-pairing chromatography negative-mode electrospray-ionization high-resolution MS on a stand-alone orbitrap (Thermo)<sup>8</sup>. Fluxes from <sup>14</sup>C-labelled substrates to CO<sub>2</sub> were measured by adding trace <sup>14</sup>C-labelled nutrient to normal culture media, quantifying radioactive CO<sub>2</sub> release, and correcting for intracellular substrate labelling according to percentage of radioactive tracer in the media and fraction of particular intracellular metabolite deriving from media uptake, as measured using <sup>13</sup>C-tracer. To assess the potential contribution of various metabolic pathways to NADPH production, we analysed feasible steady-state fluxes of a genome-scale human metabolic network model<sup>12</sup> constrained by experimentally measured uptake and excretion fluxes and growth rate of the iBMK cell line. The flux balance equations were solved in MATLAB with the objective function formulated to minimize the total sum of fluxes<sup>14</sup>. NADPH consumption by reductive biosynthesis was determined based on reaction stoichiometries, experimentally measured cellular biomass composition, growth rate, fractional *de novo* synthesis of fatty acids (by <sup>13</sup>C-labelling from [U-<sup>13</sup>C]glucose and [U-<sup>13</sup>C]glutamine), and fractional synthesis of proline from glutamate versus arginine (by <sup>13</sup>C-labelling from [U-<sup>13</sup>C]glutamine). Correction for the deuterium kinetic isotope effect was based on the assumption that total metabolic fluxes are not impacted. Let  $x$  be the fractional labelling of the relevant substrate hydrogen,  $F_U$  be the NADPH production flux from unlabelled substrate and  $F_L$  be the NADPH production flux from the labelled substrate.

$$\frac{F_L}{F_U} = \frac{x}{1-x} \left( \frac{V_H/V_D}{1-x} \right) \quad (2)$$

$$F_{\text{reaction}} = F_L + F_U = F_L \frac{(V_H/V_D) + x(1 - (V_H/V_D))}{x} \quad (3)$$

$F_L/x$  is the flux in cases without a discernible kinetic isotope effect (for example, for <sup>13</sup>C). The remaining term is the correction factor for the kinetic isotope effect:

$$C_{\text{KIE}} = \frac{V_H}{V_D} + x \left( 1 - \frac{V_H}{V_D} \right) \quad (4)$$

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 March 2013; accepted 6 March 2014.

Published online 4 May 2014.

- Voet, D. V. & Voet, J. G. *Biochemistry* 3rd edn (John Wiley & Sons, 2004).
- Jiang, P., Du, W., Mancuso, A., Wellen, K. E. & Yang, X. Reciprocal regulation of p53 and malic enzymes modulates metabolism and senescence. *Nature* **493**, 689–693 (2013).
- Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496**, 101–105 (2013).
- Lee, W. N. *et al.* Mass isotopomer study of the nonoxidative pathways of the pentose cycle with [1,2-<sup>13</sup>C<sub>2</sub>]glucose. *Am. J. Physiol.* **274**, E843–E851 (1998).

- Metallo, C. M., Walther, J. L. & Stephanopoulos, G. Evaluation of <sup>13</sup>C isotopic tracers for metabolic flux analysis in mammalian cells. *J. Biotechnol.* **144**, 167–174 (2009).
- Fan, T. W. *et al.* Rhabdomyosarcoma cells show an energy producing anabolic metabolic phenotype compared with primary myocytes. *Mol. Cancer* **7**, 79 (2008).
- Brekke, E. M., Walls, A. B., Schousboe, A., Waagepetersen, H. S. & Sonnewald, U. Quantitative importance of the pentose phosphate pathway determined by incorporation of <sup>13</sup>C from [2-<sup>13</sup>C]- and [3-<sup>13</sup>C]glucose into TCA cycle intermediates and neurotransmitter amino acids in functionally intact neurons. *J. Cereb. Blood Flow Metab.* **32**, 1788–1799 (2012).
- Lu, W. *et al.* Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. *Anal. Chem.* **82**, 3212–3221 (2010).
- Circu, M. L., Maloney, R. E. & Aw, T. Y. Disruption of pyridine nucleotide redox status during oxidative challenge at normal and low-glucose states: implications for cellular adenosine triphosphate, mitochondrial respiratory activity, and reducing capacity in colon epithelial cells. *Antioxid. Redox Signal.* **14**, 2151–2162 (2011).
- Shreve, D. S. & Levy, H. R. Kinetic mechanism of glucose-6-phosphate dehydrogenase from the lactating rat mammary gland. Implications for regulation. *J. Biol. Chem.* **255**, 2670–2677 (1980).
- Price, N. E. & Cook, P. F. Kinetic and chemical mechanisms of the sheep liver 6-phosphogluconate dehydrogenase. *Arch. Biochem. Biophys.* **336**, 215–223 (1996).
- Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl Acad. Sci. USA* **104**, 1777–1782 (2007).
- Degenhardt, K., Chen, G., Lindsten, T. & White, E. BAX and BAK mediate p53-independent suppression of tumorigenesis. *Cancer Cell* **2**, 193–203 (2002).
- Folger, O. *et al.* Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* **7**, 501 (2011).
- Tibbetts, A. S. & Appling, D. R. Compartmentalization of mammalian folate-mediated one-carbon metabolism. *Annu. Rev. Nutr.* **30**, 57–81 (2010).
- Christensen, K. E. & Mackenzie, R. E. Mitochondrial methylenetetrahydrofolate dehydrogenase, methenyltetrahydrofolate cyclohydrolase, and formyltetrahydrofolate synthetases. *Vitam. Horm.* **79**, 393–410 (2008).
- Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nature Genet.* **43**, 869–874 (2011).
- Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**, 346–350 (2011).
- Maddocks, O. D. *et al.* Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. *Nature* **493**, 542–546 (2013).
- Zhang, W. C. *et al.* Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell* **148**, 259–272 (2012).
- Nilsson, R. *et al.* Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nature Commun.* **5**, 3128 (2014).
- Ayromlou, H., Hajipour, B., Hossenian, M. M., Khodadadi, A. & Vatankhah, A. M. Oxidative effect of methotrexate administration in spinal cord of rabbits. *J. Pak. Med. Assoc.* **61**, 1096–1099 (2011).
- Bradley, K. G. & Bradley, M. E. Purine nucleoside-dependent inhibition of cellular proliferation in 1321N1 human astrocytoma cells. *J. Pharmacol. Exp. Ther.* **299**, 748–752 (2001).
- Tedeschi, P. M. *et al.* Contribution of serine, folate and glycine metabolism to the ATP, NADPH and purine requirements of cancer cells. *Cell Death Dis.* **4**, e877 (2013).
- Ye, J. *et al.* Pyruvate kinase M2 promotes *de novo* serine synthesis to sustain mTORC1 activity and cell proliferation. *Proc. Natl Acad. Sci. USA* **109**, 6904–6909 (2012).
- Chaneton, B. *et al.* Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature* **491**, 458–462 (2012).
- Anastasiou, D. *et al.* Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. *Science* **334**, 1278–1283 (2011).
- Jain, M. *et al.* Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **336**, 1040–1044 (2012).
- Warburg, O. On the origin of cancer cells. *Science* **123**, 309–314 (1956).
- Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).

**Acknowledgements** The iBMK parental and Akt cell lines were generously provided by E. White. The <sup>14</sup>C-labelled CO<sub>2</sub> release experiments were conducted with the help of E. Suh and H. Collier. NMR measurement of formate was carried out with the help of I. Lewis. We thank H. Djabballah and the High-Throughput Drug Screening Facility at MSKCC for supplying the hairpins, and M. Vander Heiden and his laboratory members for discussions. This work was supported by Stand Up To Cancer and NIH R01 grants CA163591, AI097382, and CA105463, P01 grant CA104838 and P50 grant GM071508. J.F. is a Howard Hughes Medical Institute (HHMI) international student research fellow. J.J.K. is a Hope Funds for Cancer Research fellow (HFRC-11-03-01).

**Author Contributions** J.F. and J.D.R. conceived the study. J.F., J.Y., C.B.T. and J.D.R. designed the experiments. J.F., J.Y. and J.J.K. performed the experiments. T.S. and J.F. conducted the computational analyses. J.D.R. and J.F., assisted by J.Y., T.S. and C.B.T., wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Readers are welcome to comment on the online version of the paper. The authors declare competing financial interests: details are available in the online version of the paper. Correspondence and requests for materials should be addressed to J.D.R. ([josh@genomics.princeton.edu](mailto:josh@genomics.princeton.edu)).

## Methods

**Cell lines and culture conditions.** HEK293T (large T antigen-transformed human embryonic kidney cells) and MDA-MB-468 (triple-negative human breast cancer cells) were purchased from ATCC. Immortalized baby mouse kidney epithelial cells (iBMK) with and without myr-AKT were a gift of Eileen White<sup>13,31</sup>. All cell lines were grown in Dulbecco's modified eagle medium (DMEM) without pyruvate (CELLGRO), supplemented with 10% dialysed fetal bovine serum (Invitrogen) in a 5% CO<sub>2</sub> incubator at 37 °C.

Knockdown of enzymes were by infection with lentivirus expressing the corresponding shRNA: shMTHFD1, #1: CCGGGCTGAAGAGATTGGGATCAAACCTCGAGTTTGATCCCAATCTCTTCAGCTTTTGTG, #2: CCGGGCCATTGATGCTCGGATATTTCTCGAGAAATATCCGAGCATCAATGGCTTTTGTG; shMTHFD2, #1: CCGGGCAGTTGAAGAAACATACAATCTCGAGATTGTATGTTTCTTCAA CTGCTTTTGTG, #2: CCGGGCTGGGTATATCACTCCAGTTCTCGAGAACTG GAGTGATATACCCAGCTTTTGTG; shG6PD, #1: CCGGCAACAGATACAAGA ACGTGAACGAGTTCACGTTCTGTATCTGTTGTTTGTG, #3: CCGGGC TGATGAAGAGAGTGGGTTTCTCGAGAAACCACTCTCTTCATCAGCTTT TGTG; shNNT: CCGGCCCTATGGTTAATCCAACATTCTCGAGAAATGTTGGA TTAACCATAGGGTTTGTG; shME1, #1: CCGGGCCTTCAATGAACGGCCTA TTCTCGAGAAATAGCCGCTTCATTGAAGGCTTTTGTG, #2: CCGGGCAACAA TATAGTTTGGTGTCTCGAGAAACACCAACTATATTGTTGGTGTGTTTGTG and puromycin selection. To obtain the shRNA-expressing virus, pLKO-shRNA vectors (Sigma-Aldrich) were cotransfected with the third generation lentivirus packaging plasmids (pMDLg, pCMV-VSV-G and pRsv-Rev) into HEK293T cells using FuGENE 6 Transfection Reagent (Promega), fresh media added after 24 h, and viral supernatants collected at 48 h. Target cells were infected by viral supernatant (diluted 1:1 with DMEM; 6 µg ml<sup>-1</sup> polybrene), fresh DMEM added after 24 h, and selection with 3 µg ml<sup>-1</sup> puromycin initiated at 48 h and allowed to proceed for 2 to 3 days. Thereafter, cells were maintained in DMEM with 1 µg ml<sup>-1</sup> puromycin. For *IDH1*, *IDH2* and *ALDH1L2* knockdown, siRNA targeting *IDH1* or *IDH2* (Thermo Scientific, 40 nM) or *ALDH1L2* (Santa Cruz, 30 nM) were transfected into H293T cells using Lipofectamine RNAiMAX (Invitrogen). Knockdown of the enzymes was confirmed by immunoblotting with commercial antibodies: G6PD (Bethyl Laboratories), MTHFD1 and MTHFD2 (Abgent), IDH1 (Proteintech Group), IDH2 (Abcam) and ALDH1L2 (Santa Cruz) or quantitative RT-PCR probes (*ME1* and *NNT*, Applied Biosystems) (Extended Data Fig. 10). For enzymes with more than one successful knockdown sequence, data presented here are mean ± s.d. of independent experiments using different shRNA sequences.

**Measurement of metabolite concentrations and labelling patterns.** Cells were collected at 80% confluency. For metabolomic experiments, medium was replaced every 2 days and additionally 2 h before metabolome collection and/or isotope tracer addition. Metabolism was quenched and metabolites extracted by aspirating media and immediately adding 80:20 methanol:water at -80 °C. Supernatants from two rounds of methanol:water extraction were combined, dried under N<sub>2</sub>, resuspended in HPLC water, placed in a 4 °C autosampler, and analysed within 6 h to avoid NADPH degradation.

The LC-MS method involved reversed-phase ion-pairing chromatography coupled by negative mode electrospray ionization to a stand-alone orbitrap mass spectrometer (Thermo Scientific) scanning from *m/z* 85–1,000 at 1 Hz at 100,000 resolution<sup>8,32,33</sup> with LC separation on a Synergy Hydro-RP column (100 mm × 2 mm, 2.5 µm particle size, Phenomenex, Torrance, CA) using a gradient of solvent A (97%:3% H<sub>2</sub>O:MeOH with 10 mM tributylamine and 15 mM acetic acid), and solvent B (100% MeOH). The gradient was 0 min, 0% B; 2.5 min, 0% B; 5 min, 20% B; 7.5 min, 20% B; 13 min, 55% B; 15.5 min, 95% B; 18.5 min, 95% B; 19 min, 0% B; 25 min, 0% B. Injection volume was 10 µl, flow rate 200 µl min<sup>-1</sup>, and column temperature 25 °C. Data were analysed using the MAVEN software suite<sup>34</sup>. Data from <sup>13</sup>C-labelling experiments were adjusted for natural <sup>13</sup>C abundance and impurity of labelled substrate; those from <sup>2</sup>H-labelling were not adjusted (natural <sup>2</sup>H abundance is negligible)<sup>35</sup>. The absolute concentration of 6-phosphogluconate was quantified by comparing the signal of <sup>13</sup>C-labelled intracellular compound (from feeding [U-<sup>13</sup>C]glucose) to the signal of unlabelled internal standard.

**Fractional labelling of NADPH redox active site.** The fractional NADPH redox active site labelling (*x*) was measured from the observed NADPH and NADP<sup>+</sup> labelling patterns from the same sample. We calculated *x* to best fit the steady-state mass distribution vectors of NADPH and NADP<sup>+</sup> (*M*<sub>NADPH</sub> and *M*<sub>NADP<sup>+</sup></sub>) by least square fitting in MATLAB (function: lsqcurvefit).

$$M_{NADP^+} = \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ \vdots \\ m_N \end{bmatrix} \begin{matrix} M+0 \\ M+1 \\ M+2 \\ \\ M+N \end{matrix}$$

$$M_{NADPH} = \begin{bmatrix} m_0 \times (1-x) \\ m_1 \times (1-x) + m_0 \times x \\ m_2 \times (1-x) + m_1 \times x \\ \vdots \\ m_N \times (1-x) + m_{N-1} \times x \\ m_N \times x \end{bmatrix} \begin{matrix} M+0 \\ M+1 \\ M+2 \\ \\ M+N \\ M+N+1 \end{matrix} \quad (5)$$

**Network analysis of potential NADPH producing pathways.** To assess the potential contribution of various metabolic pathways to NADPH production, we analysed feasible steady-state fluxes of a genome-scale human metabolic network model<sup>12</sup>. The glucose (98 nmol/(µl × h)), glutamine (40 nmol/(µl × h)), and oxygen uptake rates (21 nmol/(µl × h)); and lactate (143 nmol/(µl × h)), alanine (2 nmol/(µl × h)), pyruvate (15 nmol/(µl × h)), and formate (< 0.25 nmol/(µl × h)) excretion rates were set to experimental measured fluxes in the iBMK cell line, as measured by a combination of electrochemistry (glucose, glutamine, lactate on YSI7200 instrument, YSI, Yellow Springs, OH), LC-MS (alanine, pyruvate with isotopic internal standards), fluorometry (oxygen on XF24 flux analyser, Seahorse Bioscience, North Billerica, MA), and NMR (formate by <sup>1</sup>H 500 MHz, Bruker, 10 µM limit of detection). The uptake of amino acids from DMEM media were bounded to not more than a third of that of glutamine, which is a loose constraint relative to experimental observations in iBMK cells and in NCI-60 cells<sup>28</sup>. Biomass requirements were based on the experimentally determined growth rate of the iBMK cell-line with protein, fatty acids and nucleotides accounting for 60%, 10% and 10% of the total cellular dry mass, respectively, based on experimental measurements. Steady-state intracellular fluxes that best fit these experimental constraints were then selected by solving the flux balance equations in MATLAB with the objective function formulated to minimize the sum of total fluxes<sup>14</sup>.

**Correction for deuterium's kinetic isotope effect.** The kinetic isotope effect (*V<sub>H</sub>*/*V<sub>D</sub>*) for isolated NADPH producing enzymes ranges from 1.8 to 4, with isolated G6PD and 6-phosphogluconate dehydrogenase having *V<sub>H</sub>*/*V<sub>D</sub>* = 1.8 (refs 10, 11). However, cellular homeostatic mechanisms (including flux control being distributed across multiple pathway enzymes) may result in a lesser impact on labelling patterns in cells.

The smallest reasonable correction for the deuterium kinetic isotope effect is based on the assumption that total metabolic fluxes are not affected. This correction was used as the default in this work. Let *x* be the fractional labelling of the relevant substrate hydrogen, *F<sub>U</sub>* be the NADPH production flux from unlabelled substrate and *F<sub>L</sub>* be the NADPH production flux from the labelled substrate.

$$\frac{F_L}{F_U} = \frac{x}{\frac{V_H}{V_D} - 1 - x} \quad (2)$$

$$F_{\text{reaction}} = F_L + F_U = F_U \frac{(V_H/V_D) + x(1 - (V_H/V_D))}{x} \quad (3)$$

*F<sub>L</sub>*/*x* is the flux in cases without a discernible kinetic isotope effect (for example, for <sup>13</sup>C). The remaining term is the correction factor for the kinetic isotope effect:

$$C_{\text{KIE}} = \frac{V_H}{V_D} + x \left( 1 - \frac{V_H}{V_D} \right) \quad (4)$$

The largest reasonable correction for the deuterium kinetic isotope effect is based on the assumption that pathway flux is decreased by the introduction of <sup>2</sup>H-labelled tracer equivalent to the decrease in activity of the associated enzyme observed *in vitro*:

$$C_{\text{KIE}} = \frac{V_H/V_D}{1 + N \times ((V_H/V_D) - 1) \times X_{\text{NADPH}}} \quad (6)$$

in which *N* is the number of NADPH produced per substrate molecule passing through the pathway. For the oxidative pentose phosphate pathway, *N* = 2. Note that the effect of the kinetic isotope effect on [<sup>2</sup>H]NADPH production may be partially offset by an analogous (albeit smaller) kinetic isotope effect in [<sup>2</sup>H]NADPH consuming reactions. *V<sub>H</sub>*/*V<sub>D</sub>* for fatty acid synthetase is ~1.1 (ref. 36). The effect of different mechanisms of correcting for the deuterium kinetic isotope is shown in Extended Data Fig. 1.

**Quantifying absolute oxidative pentose phosphate pathway flux based on 6-phosphogluconate labelling kinetics.** To quantify the absolute oxidative pentose phosphate pathway flux, cells were switched to media containing [U-<sup>13</sup>C]glucose, and the kinetics glucose-6-phosphate and 6-phosphogluconate labelling were measured. The unlabelled fraction of 6-phosphogluconate decays with time as:



$$\frac{d[6\text{-phosphogluconate}]^{\text{unlabelled}}}{dt} = -F_{\text{oxPPP}} \frac{[6\text{-phosphogluconate}]^{\text{unlabelled}}}{[6\text{-phosphogluconate}]^{\text{total}}} + F_{\text{oxPPP}} \times \text{Fraction}^{\text{unlabelledG6P}}(t) \quad (7)$$

where  $F_{\text{oxidative pentose phosphate pathway}}$  ( $F_{\text{oxPPP}}$ ) is the flux of oxidative pentose phosphate pathway,  $[6\text{-phosphogluconate}]^{\text{total}}$  is the total cellular 6-phosphogluconate concentration, which was directly measured, and  $\text{Fraction}^{\text{unlabelledG6P}}(t)$  is the unlabelled fraction of glucose-6-phosphate at time  $t$ , which decays exponentially.  $F_{\text{oxPPP}}$  was obtained by least square fitting as per ref. 37.

**Quantifying the upper limit of NADPH production via malic enzyme by  $^{13}\text{C}$  labelling.** Malic enzyme (ME) can produce either NADH or NADPH. Thus, total malic enzyme flux puts an upper limit on the associated NADPH production. To probe overall malic enzyme activity, cells were incubated with  $[\text{U-}^{13}\text{C}]\text{glutamine}$  for 48 h, which resulted in the majority of intracellular malate being uniformly labelled ( $^{13}\text{C}_4$ ), with a small portion being  $^{13}\text{C}_3$ . For simplicity, we assume that  $^{13}\text{C}_3$ -malate is an equal mix of  $[1,2,3\text{-}^{13}\text{C}_3]\text{malate}$  and  $[2,3,4\text{-}^{13}\text{C}_3]\text{malate}$  owing to rapid interconversion with fumarate (which is symmetric). Malic enzyme produces  $^{13}\text{C}_3$  pyruvate from both  $^{13}\text{C}_4$  malate and  $[1,2,3\text{-}^{13}\text{C}_3]\text{malate}$ , whereas glycolysis produces unlabelled pyruvate (See Extended Data Fig. 4).

$$\text{Flux}_{\text{NADPH ME}} \leq \frac{[^{13}\text{C}_3]\text{Pyruvate}}{\text{Total pyruvate}} \times \frac{\text{Total malate}}{[^{13}\text{C}_4]\text{Malate} + 0.5[^{13}\text{C}_3]\text{Malate}} \times \text{Flux}_{\text{glycolysis}} \quad (8)$$

**Estimation of fractional contribution of MTHFD to NADPH production based on  $^2\text{H}$ -serine labelling.** Similar to quantifying relative contribution of oxidative pentose phosphate pathway to cytosolic NADPH production, the contribution of THF-pathway can be estimated from  $^2\text{H}$ -serine labelling as follows:

$$\text{Fraction}_{\text{NADPH THF-pathway}} = \frac{[^2\text{H}]\text{NADPH}}{\text{Total } [^2\text{H}]\text{NADPH}} \times \frac{\text{Total methylene-THF}}{[^2\text{H}]\text{methylene-THF}} \times C_{\text{KIE}}(\text{MTHFD}) \quad (9)$$

Existing methods do not allow direct measurement of methylene-THF labelling, but such labelling can be approximated based on intracellular serine labelling (formally, the  $^2\text{H}$ -serine labelling places an upper bound on  $^2\text{H}$ -methylene-THF labelling).

$$\text{Fraction}_{\text{NADPH THF-pathway}} \geq \frac{[^2\text{H}]\text{NADPH}}{\text{Total } [^2\text{H}]\text{NADPH}} \times \frac{\text{Total serine}}{[^2\text{H}]\text{serine}} \times C_{\text{KIE}}(\text{MTHFD}) \quad (10)$$

MTHFD1 has a deuterium kinetic isotope effect  $V_{\text{H}}/V_{\text{D}}$  of  $\sim 3$ .

**Measurement of  $^{14}\text{CO}_2$  release.** Radioactive  $\text{CO}_2$  released by cells from positionally labelled substrates was measured by trapping the  $\text{CO}_2$  in filter paper saturated with 10 M KOH as previously described<sup>14</sup>. Cells were grown in tissue culture flasks with DMEM medium with less than normal bicarbonate (0.74 g per l) and addition of HEPES buffer (6 g per l, pH 7.4). At the beginning of experiment, trace amount of desired  $^{14}\text{C}$ -labelled tracer was added to the media. For each cell line, the amount was selected to be the minimum that gives a sufficient radioactive  $\text{CO}_2$  signal to quantitate accurately ( $\sim 1 \mu\text{Ci ml}^{-1}$ ). All knockdown lines were treated identically to their corresponding parental line. Then the flask was sealed with a rubber stopper with a central well (Kimble Chase) containing a piece of filter paper saturated with 10 M KOH solution. The flasks were incubated at  $37^\circ\text{C}$  for 24 h.  $\text{CO}_2$  released by cells was absorbed by the base (that is, KOH) in the central well. Metabolism was stopped by injection of 1 ml 3 M acetic acid solution through the rubber stopper. The flasks were then incubated at room temperature for 1 h to ensure all the  $\text{CO}_2$  dissolved in media was released and absorbed into the central well. The filter paper and all the liquid in central well was transfer to a scintillation vial containing 15 ml liquid scintillation cocktail (PerkinElmer). The central well was washed with 100  $\mu\text{l}$  of water twice, and the water was added to the same scintillation vial. Radioactivity was measured by liquid scintillation counting. In parallel, the same experiments were performed using  $[\text{U-}^{13}\text{C}]$ -labelled nutrient (in amounts that fully replaced the unlabelled nutrient in DMEM) and the extent of labelling of the intracellular metabolite, that is the substrate of the  $\text{CO}_2$ -releasing reaction, was measured by LC-MS. Absolute  $\text{CO}_2$  release rates from the nutrients of interest were calculated as follows:

$$\text{Rate}_{\text{CO}_2 \text{ from source } i} [\text{nmol/h}/\mu\text{l cells}] = \frac{\text{Rate}_{\text{CO}_2 \text{ from } ^{14}\text{C-labelled tracer}_i} [\mu\text{Ci/h}/\mu\text{l cells}]}{\text{Overall media tracer}_i \text{ activity} [\mu\text{Ci/nmol}]} \times \frac{1}{\text{Fraction}_{\text{intracellular compound}_i \text{ from media}}} \quad (11)$$

**Fractional labelling of cytosolic formyl groups from  $[\text{U-}^{13}\text{C}]\text{serine}$ .** Cells were cultured with media containing  $[\text{U-}^{13}\text{C}]\text{serine}$  for 48 h, washed three times with cold PBS to remove extracellular serine, extracted, and the intracellular labelling pattern analysed by LC-MS for ATP (representing purines; there is no labelling of ribose-phosphate based on LC-MS measurements), glycine and serine. The purine ring has 5 carbons: 1 from  $\text{CO}_2$ , 2 from glycine and 2 from formyl groups (from 10-formyl-THF). Assuming that  $\text{CO}_2$  labelling is negligible, which is realistic for cells grown in a 5%  $\text{CO}_2$  incubator, let  $X_{\text{ATP-}i}$  and  $X_{\text{Gly-}j}$  represent the experimentally observed fraction ATP and glycine with  $i$  and  $j$  labelled carbons. The cytosolic 10-formyl-THF labelling fraction,  $y$ , was fit by least squares

$$\begin{aligned} X_{\text{ATP-0}} &= X_{\text{Gly-0}} \times (1-y)^2 \\ X_{\text{ATP-1}} &= 2 \times X_{\text{Gly-0}} \times y(1-y) \\ X_{\text{ATP-2}} &= X_{\text{Gly-2}} \times (1-y)^2 + X_{\text{Gly-0}} \times y^2 \\ X_{\text{ATP-3}} &= 2 \times X_{\text{Gly-2}} \times y(1-y) \\ X_{\text{ATP-4}} &= X_{\text{Gly-2}} \times y^2 \end{aligned} \quad (12)$$

**Cytosolic NADPH production from 10-formyl-THF pathway.** Cytosolic NADPH production from 10-formyl-THF pathway was quantified by tracking its end products: 10-formyl-THF consumed by purine synthesis and  $\text{CO}_2$ . (Formate excretion into media is below the detection limit of NMR.) All 10-formyl-THF consumed by purine synthesis is generated in cytosol and associated with production of 1 NADPH. For each  $\text{CO}_2$  released from serine C3, if the reaction happens in cytosol, 1 NADPH is produced from 10-formyl-THF oxidation, and another NADPH is produced via MTHFD1. Total cytosolic NADPH production via 10-formyl-THF pathway is:

$$\text{Flux}_{\text{NADPH from THF pathway}} = 2 \times \text{Flux}_{\text{purine synthesis}} + 2 \times \text{Flux}_{\text{CO}_2 \text{ from serine C3}} \quad (13)$$

If complete oxidation of serine C3 instead happens in mitochondria, there is no cytosolic NADPH production associated with  $\text{CO}_2$  released from serine C3 (that is, no red bar in Fig. 3d). Instead, one mitochondrial NADPH is produced from 10-formyl-THF oxidation, and zero to one other mitochondrial NADPH from 5, 10-methylene-THF oxidation depending on the enzyme used to catalyse the reaction and its cofactor specificity. In mitochondria, this reaction can be catalysed by MTHFD2, which (at least in the presence of high phosphate *in vitro*) preferentially uses  $\text{NAD}^+$  or by MTHFD2L, which uses  $\text{NADP}^+$ .

**ROS measurement, cell proliferation and cell death assay.** ROS measurement followed published protocols<sup>38</sup>. Briefly, HEK293T cells were incubated with  $5 \mu\text{M}$  CM-H2DCFDA (Invitrogen) for 30 min. Cells were trypsinized, and mean FL1 fluorescence was measured by flow cytometry. Cell proliferation was measured by trypsinizing cells and counting using a Beckman's Multisizer 4 Coulter Counter. To measure cell death, cells were stained with Trypan Blue. Stained and unstained cells were counted and cell death percentages tabulated.

**Quantification of NADPH consumption by reductive biosynthesis.** The general strategy for measuring consumption fluxes was as follows: (1) identifying the biomass components produced in cells grown in DMEM by NADPH-driven reductive biosynthesis (these are DNA, proline and fatty acids); (2) determining the biomass fraction of each component in each cell line; (3) quantifying the cellular growth rate  $R_{\text{growth}} = \ln(2)/t_{1/2}$ ; (4) measuring the fractional contribution of different biosynthetic routes to each biomass component via experiments with  $^{13}\text{C}$ -labelled glucose and/or glutamine and LC-MS analysis; (5) computing the average number of NADPH per unit of biomass component, which equals the sum of the fractional contribution of each route multiplied by the number of NADPH consumed by that route; and (6) determining NADPH consumption as follows:

$$\text{Consumption flux} = \left( \frac{\text{Product abundance}}{\text{Cell volume}} \right) \times R_{\text{growth}} \times (\text{Average NADPH per product}) \quad (14)$$

The required data were acquired as follows below.

**DNA.** Cellular DNA and RNA were extracted and separated with TRIzol reagent (Invitrogen), purified and quantified by Nanodrop spectrophotometer.

**Fatty acids.** Total cellular lipid was extracted and saponified after addition of isotope-labelled internal standards for the C16:0, C16:1, C18:0, and C18:1. Samples were analysed by negative ESI-LC-MS with LC separation on a C8 column. Concentrations of other fatty acids, for which isotope-labelled internal standard were not available, were measured by comparison to the palmitate internal standard. The calculated fatty acid concentrations were multiplied with a correction factor to account for incomplete lipid recovery in the first step of the sample preparation procedure. This correction factor was empirically determined to be 1.9 by experiments in which lipid standards were spiked into extraction solution.



The extent of fatty acid synthesis and elongation (both of which consume NADPH) was determined by feeding cells [U-<sup>13</sup>C]glucose and [U-<sup>13</sup>C]glutamine for multiple doublings to achieve pseudo-steady state labelling of their lipid pools. Fatty acid labelling patterns were measured and computationally simulated to quantify the fraction of production versus import for each individual fatty acid species. Extended Data Fig. 9 shows the associated data for C16:0, C16:1, C18:0, and C18:1, which together account for ~80% of total cellular fatty acids and >90% of non-essential fatty acids (essential fatty acids are imported, not synthesized, and thus do not affect NADPH production). NADPH calculations include similar data for all measurable fatty acids.

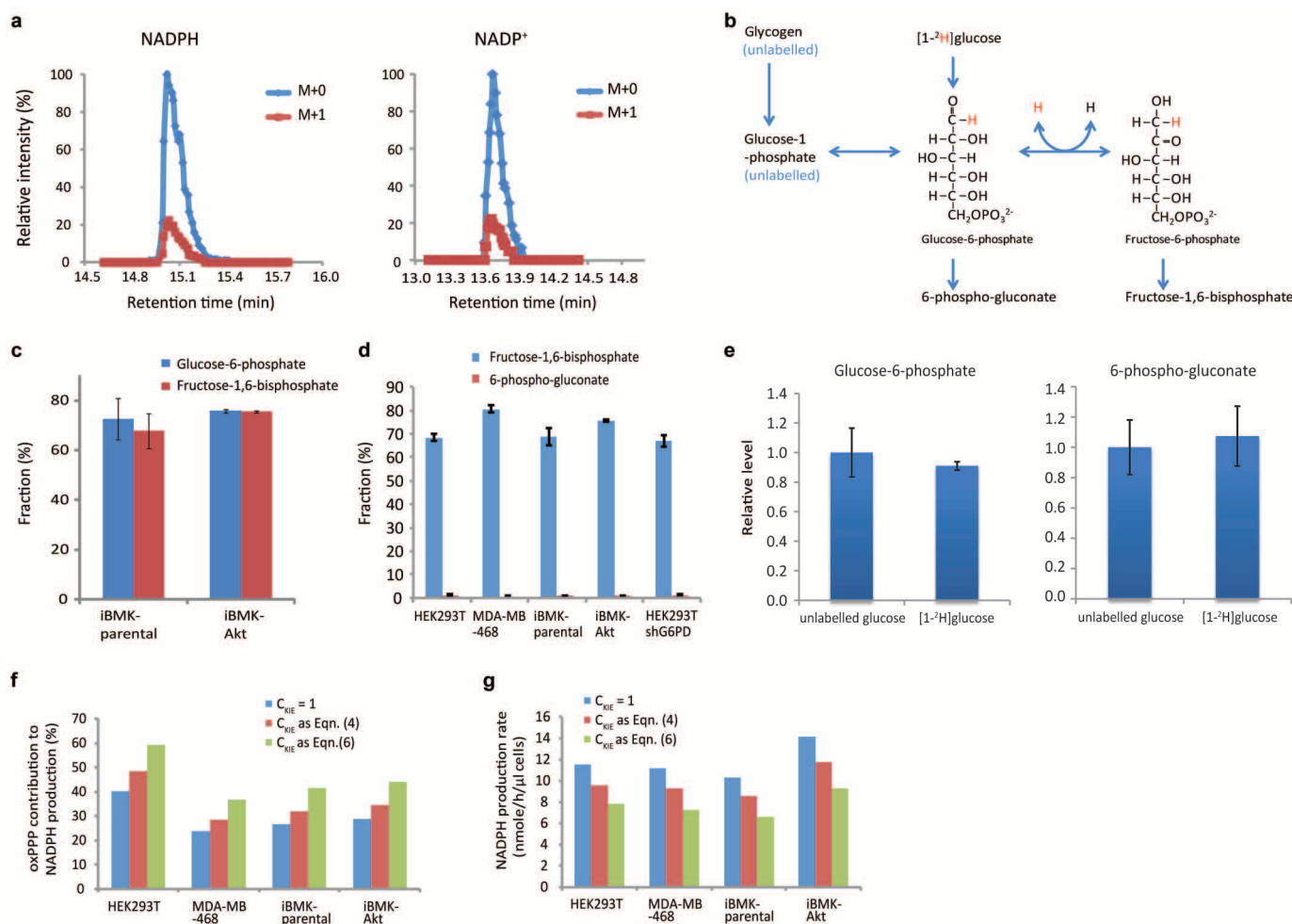
**Proline.** Proline can be made from either arginine or glutamate. Proline synthesis from either substrate requires two high-energy electrons at the step catalysed by pyrroline-5-carboxylate reductase, which may use NADH or NADPH (for simplicity, we assume an equally contribution from each). Proline synthesis from glutamate consumes one additional NADPH<sup>39</sup>. To quantify the fraction of proline synthesized from each substrate, cells were labelled with [U-<sup>13</sup>C]glutamine to steady state, which labels glutamate but not arginine. Labelling of intracellular proline and glutamate were measured.

$$X_{\text{Glu}} = \frac{\text{Fraction proline } ^{13}\text{C-labelled}}{\text{Fraction glutamate } ^{13}\text{C-labelled}} \quad (15)$$

$$\text{Flux}_{\text{NADPH for proline}} = \frac{\text{Growth rate} \times \text{Protein content}}{\text{Average formula weight per residue}} \times \text{Proline frequency} \times (1.5X_{\text{Glu}} + 0.5(1 - X_{\text{Glu}})) \quad (16)$$

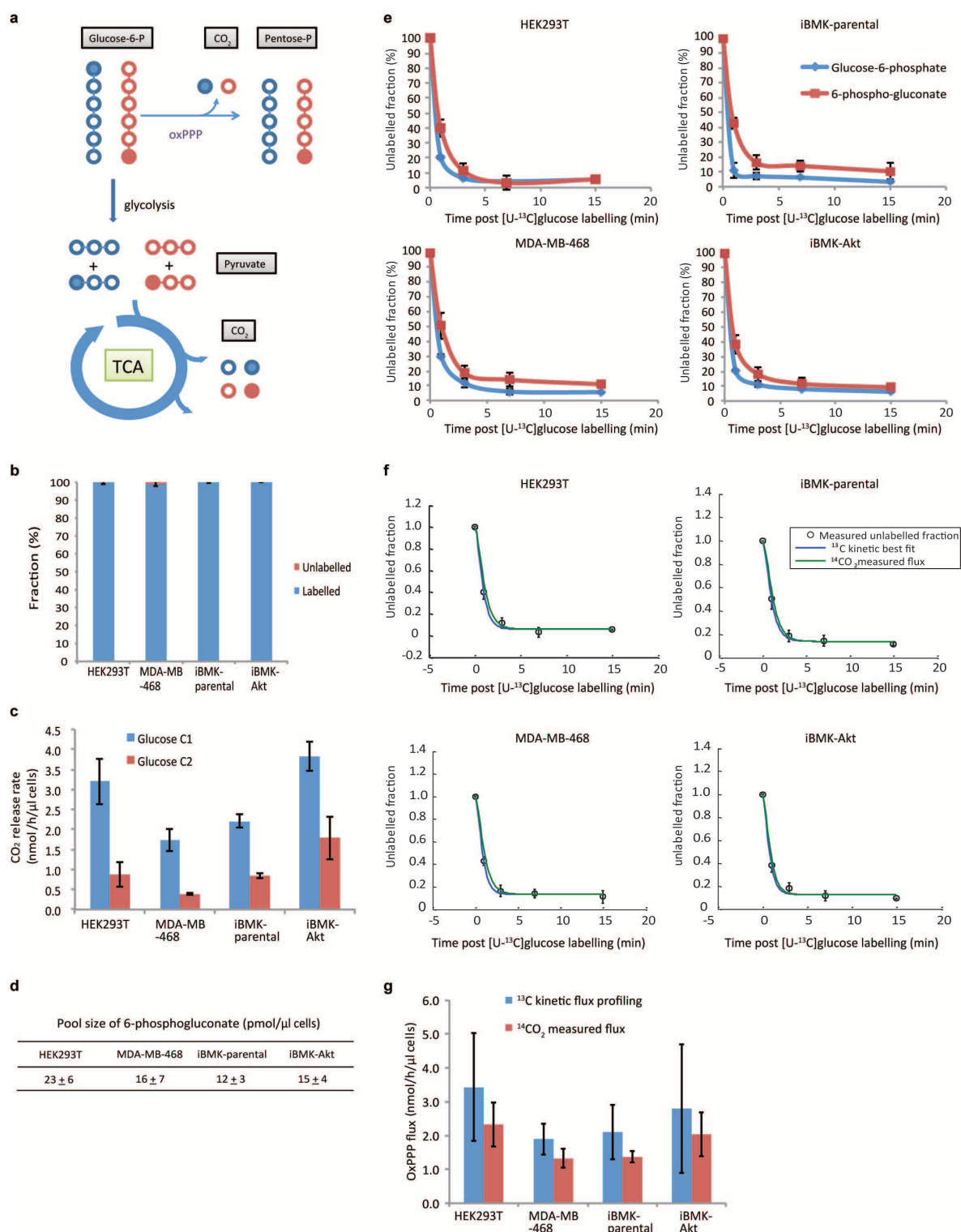
Proline synthesis enzymes are present in both the cytosol and mitochondria. For simplicity, Fig. 4 assumes exclusively cytosolic proline synthesis.

31. Mathew, R., Degenhardt, K., Haramaty, L., Karp, C. M. & White, E. Immortalized mouse epithelial cell models to study the role of apoptosis in cancer. *Methods Enzymol.* **446**, 77–106 (2008).
32. Munger, J. *et al.* Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nature Biotechnol.* **26**, 1179–1186 (2008).
33. Lemons, J. M. *et al.* Quiescent fibroblasts exhibit high metabolic activity. *PLoS Biol.* **8**, e1000514 (2010).
34. Melamud, E., Vastag, L. & Rabinowitz, J. D. Metabolomic analysis and visualization engine for LC-MS data. *Anal. Chem.* **82**, 9818–9826 (2010).
35. Millard, P., Letisse, F., Sokol, S. & Portais, J. C. IsoCor: correcting MS data in isotope labeling experiments. *Bioinformatics* **28**, 1294–1296 (2012).
36. Yuan, Z. & Hammes, G. G. Elementary steps in the reaction mechanism of chicken liver fatty acid synthase. pH dependence of NADPH binding and isotope rate effect for beta-ketoacyl reductase. *J. Biol. Chem.* **259**, 6748–6751 (1984).
37. Yuan, J., Bennett, B. D. & Rabinowitz, J. D. Kinetic flux profiling for quantitation of cellular metabolic fluxes. *Nature Protocols* **3**, 1328–1340 (2008).
38. Eruslanov, E. & Kusmartsev, S. Identification of ROS using oxidized DCFDA and flow-cytometry. *Methods Mol. Biol.* **594**, 57–72 (2010).
39. Lorans, G. & Phang, J. M. Proline synthesis and redox regulation: differential functions of pyrroline-5-carboxylate reductase in human lymphoblastoid cell lines. *Biochem. Biophys. Res. Commun.* **101**, 1018–1025 (1981).
40. Pawelek, P. D. & MacKenzie, R. E. Methenyltetrahydrofolate cyclohydrolase is rate limiting for the enzymatic conversion of 10-formyltetrahydrofolate to 5,10-methylenetetrahydrofolate in bifunctional dehydrogenase-cyclohydrolase enzymes. *Biochemistry* **37**, 1109–1115 (1998).



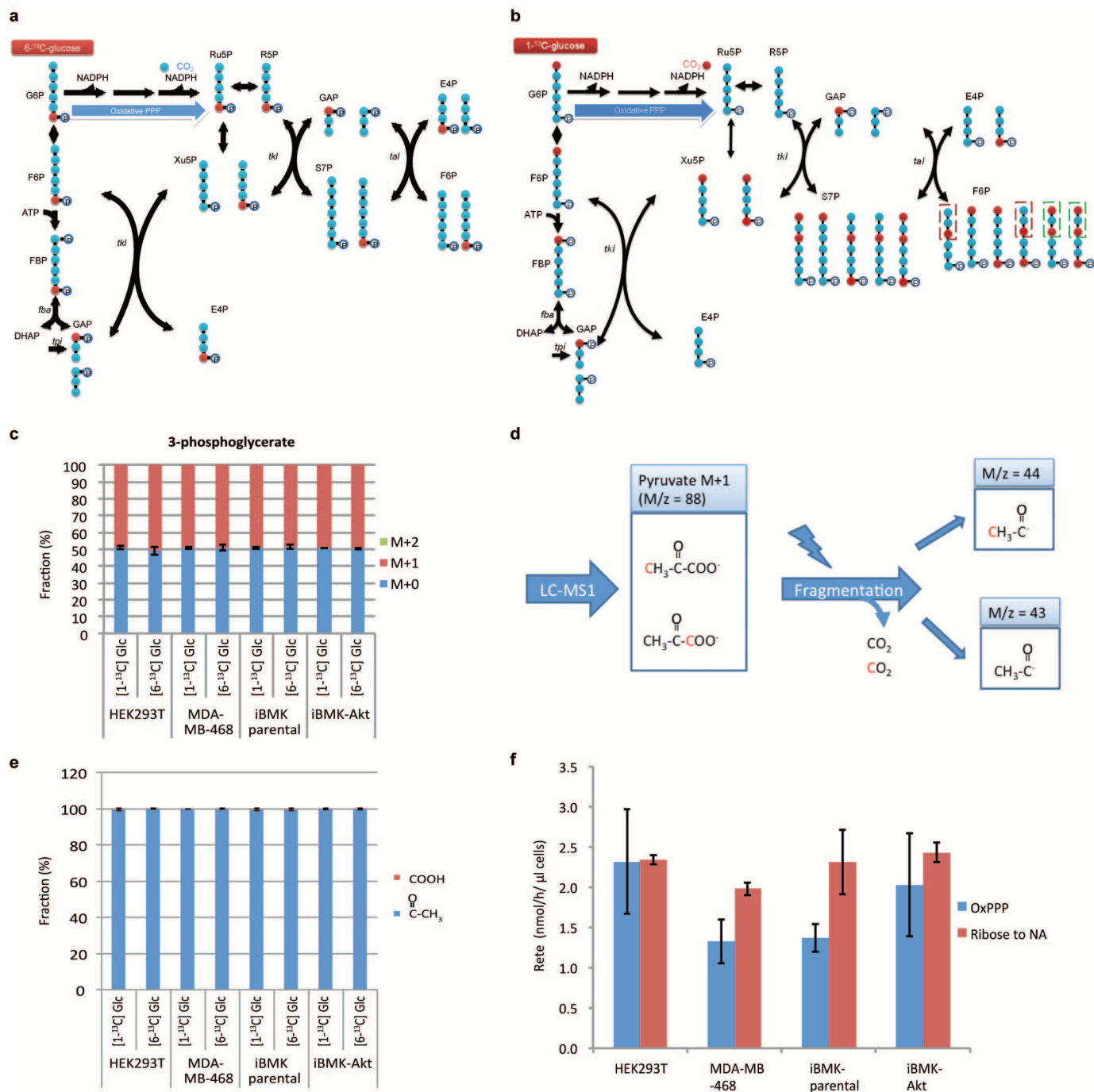
**Extended Data Figure 1 | Probing the fractional contribution of the oxidative pentose phosphate pathway to NADPH production with [<sup>2</sup>H]glucose.** **a**, Example of LC-MS chromatogram of M+0 and M+1 forms of NADPH and NADP<sup>+</sup>. Plotted values are 5 p.p.m. mass window around each compound. **b**, Extent of NADPH labelling must be corrected for extent of glucose-6-phosphate labelling. Incomplete labelling can occur due to influx from glycogen or hydrogen-deuterium exchange. **c**, Labelling fraction of glucose-6-phosphate and fructose-1,6-phosphate in iBMK cells with and without activated Akt (20 min after switching into [1-<sup>2</sup>H]glucose). **d**, Labelling fraction of fructose-1,6-phosphate and 6-phosphogluconate after feeding [1-<sup>2</sup>H]glucose. Labelling fraction of fructose-1,6-phosphate reflects the labelling of glucose-6-phosphate, whose peak after addition of the [1-<sup>2</sup>H]glucose was not sufficiently resolved from other LC-MS peaks in HEK293T and MDA-MB-468 cells to allow precise quantification of its labelling directly. The difference in the labelling fraction between glucose-6-phosphate and 6-phosphogluconate reflects the fraction of deuterium labelling specifically at position 1 of glucose-6-phosphate. **e**, Due to the kinetic isotope effect, feeding of

deuterium tracer can potentially alter pathway fluxes. To assess whether the feeding of [1-<sup>2</sup>H]glucose creates a bottleneck in the oxidative pentose phosphate pathway, we measured the relative concentration of oxidative pentose phosphate pathway intermediates with or without feeding of [1-<sup>2</sup>H]glucose. No significant changes were observed. **f**, Effect of different mechanisms of correcting for the deuterium kinetic isotope effect on fractional contribution of oxidative pentose phosphate pathway to NADPH production. **g**, Effect of different mechanisms of correcting for the deuterium kinetic isotope effect on calculated total NADPH production rate. The correction mechanisms are: (1) no kinetic isotope effect (C<sub>KIE</sub> = 1), (2) no effect on total pathway flux but preferential utilization of <sup>1</sup>H over <sup>2</sup>H-labelled substrate (equation (4) of main text) (the smallest reasonable correction, and the one applied in the main text), or (3) full kinetic isotope effect observed for the isolate enzyme with associated decrease in total pathway flux (Eqn. 6 of Methods) (the largest reasonable correction). All results are mean ± s.d., *n* ≥ 2 biological replicates from a single experiment and results were confirmed in multiple experiments.



**Extended Data Figure 2 | Two independent measurement methods give consistent oxidative pentose phosphate pathway fluxes.** **a**, Diagram of [<sup>1-<sup>14</sup>C</sup>]glucose and [<sup>6-<sup>14</sup>C</sup>]glucose metabolism through glycolysis and the oxidative pentose phosphate pathway. The oxidative pentose phosphate pathway specifically releases glucose C1 as CO<sub>2</sub>, whereas all other CO<sub>2</sub>-releasing reactions are downstream of triose phosphate isomerase (TPI). As TPI renders C1 and C6 of glucose indistinguishable (both positions become C3 of glyceraldehyde-3-phosphate), the difference in CO<sub>2</sub> release from C1 versus C6, multiplied by two, gives the absolute rate of NADPH production via oxidative pentose phosphate pathway. A potential complication involves carbon scrambling via the reactions of the non-oxidative pentose phosphate pathway, but this was negligible (see Extended Data Fig. 3). **b**, Complete carbon

labelling of glucose-6-phosphate. Glucose-6-phosphate was labelled completely (> 99%) within 2 h of switching cells into [U-<sup>13</sup>C]glucose. **c**, CO<sub>2</sub> release rate from [<sup>1-<sup>14</sup>C</sup>]glucose and [<sup>6-<sup>14</sup>C</sup>]glucose. **d**, Pool size of 6-phosphogluconate. **e**, Kinetics of glucose-6-phosphate and 6-phosphogluconate labelling upon switching cells to [U-<sup>13</sup>C]glucose. **f**, Overlay upon the 6-phosphogluconate data from **e** of simulated labelling curves based on the flux that best fits the labelling kinetics (blue) (see Methods), and the flux from <sup>14</sup>CO<sub>2</sub> release measurements (green). **g**, Calculated fluxes and 95% confidence intervals based on kinetics of 6-phosphogluconate labelling from [U-<sup>13</sup>C]-glucose, compared to radioactive CO<sub>2</sub> release from [<sup>1-<sup>14</sup>C</sup>]glucose and [<sup>6-<sup>14</sup>C</sup>]glucose. The two approaches give consistent results, with the <sup>14</sup>CO<sub>2</sub> release data being more precise. Mean ± s.d., *n* = 3.

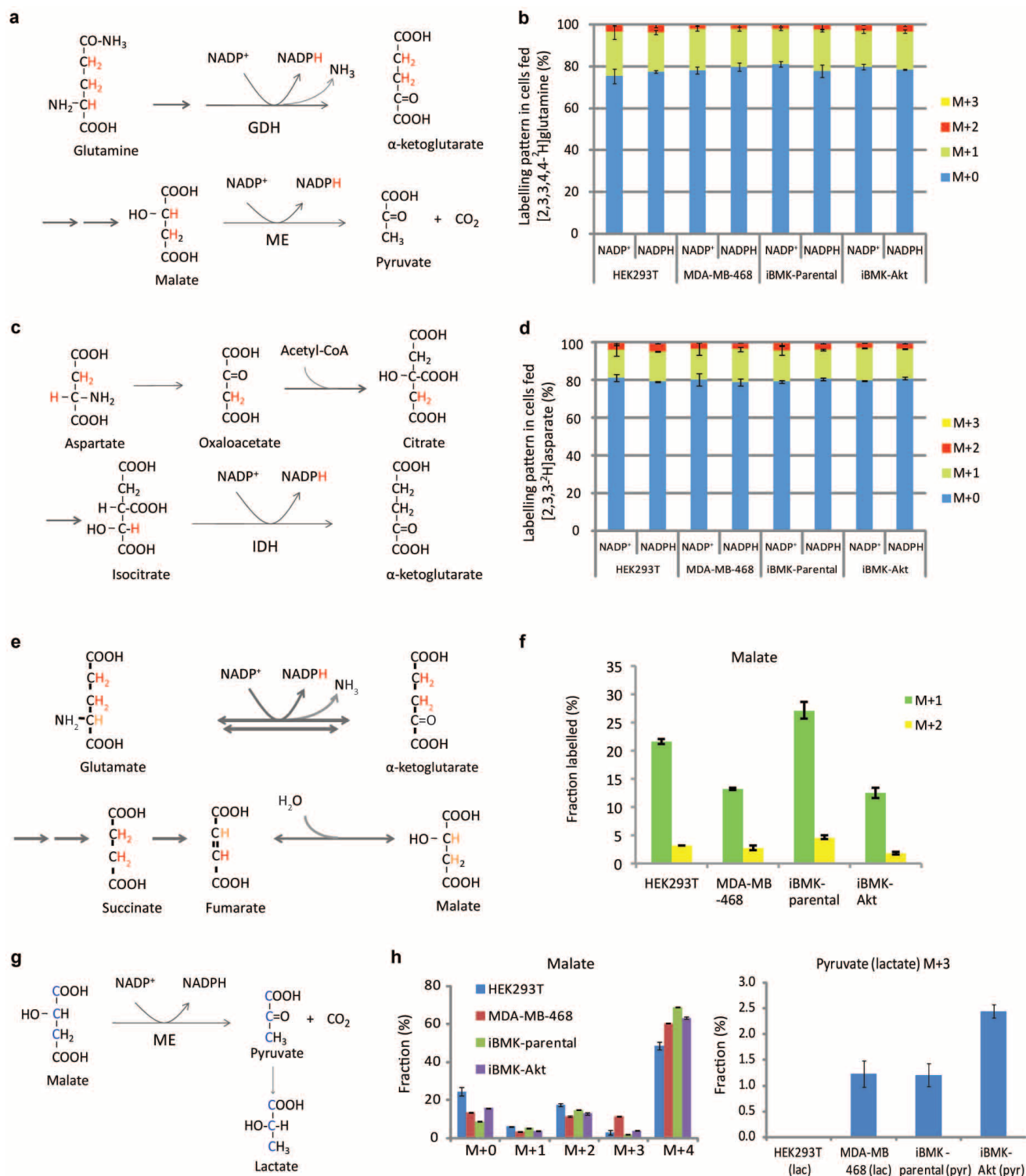


**Extended Data Figure 3 | The extent of carbon scrambling via non-oxidative pentose phosphate pathway is insufficient to substantially affect oxidative pentose phosphate pathway flux determination using [1-<sup>14</sup>C]glucose and [6-<sup>14</sup>C]glucose, with most carbon entering oxidative pentose phosphate pathway directed towards nucleotide synthesis.**

**a**, Schematic of glycolysis and pentose phosphate pathway showing fate of glucose C6. Note that glucose C6 occupies the phosphorylated position (that is, the last carbon) in every intermediate. Thus, upon catabolism to pyruvate, glucose C6 always becomes pyruvate C3, irrespective of any potential scrambling reactions. **b**, Schematic of glycolysis and pentose phosphate pathway showing fate of glucose C1. Glucose C1 can be scrambled via the non-oxidative pentose phosphate pathway, moving to C3 (red boxes) or C6 as shown here. The forms shown in the green boxes were not experimentally observed. As glucose C3 becomes pyruvate C1 (the carboxylic acid carbon of

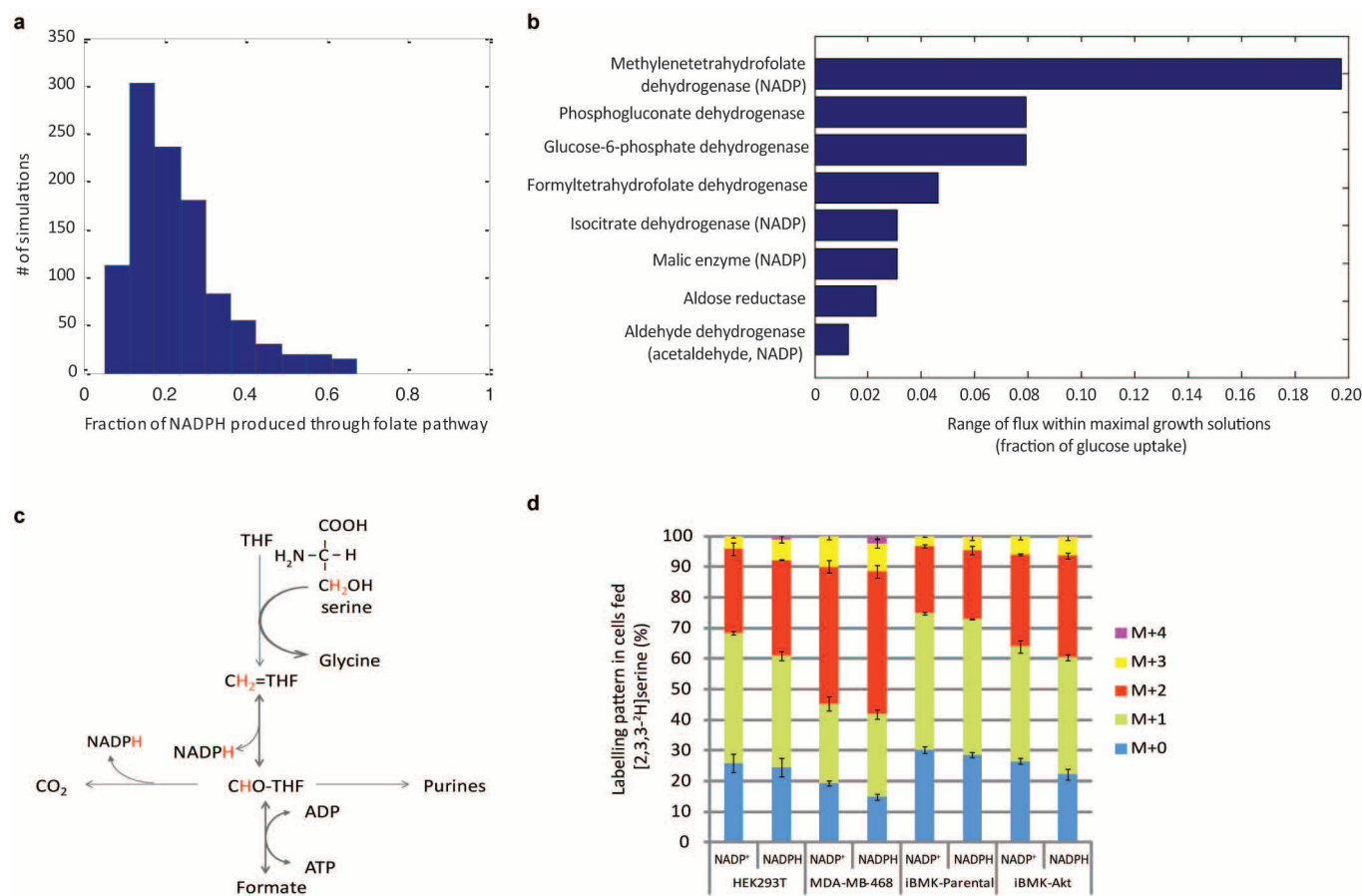
pyruvate), which is selectively released as CO<sub>2</sub> by pyruvate dehydrogenase, scrambling of C1 to C3 can potentially increase CO<sub>2</sub> release from glucose C1 relative to C6. This is ruled out in panels **d** and **e**. **c**, Feeding [1-<sup>13</sup>C]glucose or [6-<sup>13</sup>C]glucose results in 50% labelling of 3-phosphoglycerate without any double labelling (that is, M+2), as expected in the absence of scrambling. **d**, MS/MS method to analyse positional labelling of 1-labelled pyruvate. Collision induced dissociation breaks pyruvate to release the carboxylic acid group as CO<sub>2</sub>. If the daughter peak of 1-labelled pyruvate does not contain labelled carbon (*m/z* = 43), the labelling is at the C1 position; otherwise, it is at C2 or C3. **e**, After feeding [1-<sup>13</sup>C]glucose or [6-<sup>13</sup>C]glucose, pyruvate is not labelled at the C1 position (< 0.5%), ruling out extensive scrambling. **f**, Oxidative pentose phosphate pathway flux is similar to or smaller than ribose demand for nucleotide synthesis. Mean ± s.d., *n* = 3.





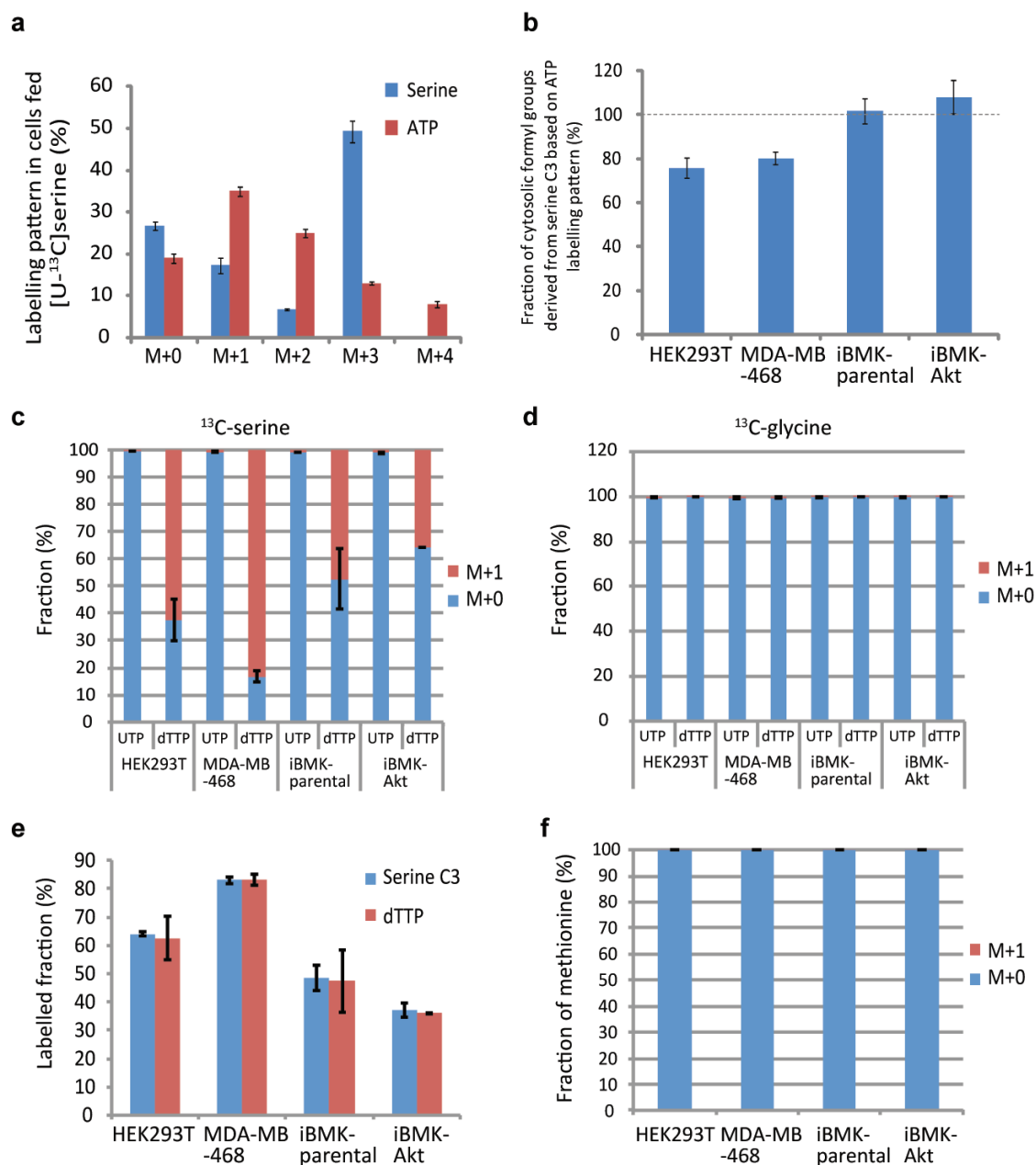
**Extended Data Figure 4 | Probing the contribution of alternative NADPH producing pathways.** **a**, Pathway diagram showing potential for [2,3,3,4,4- $^2\text{H}$ ] glutamine to label NADPH via glutamate dehydrogenase and via malic enzyme. Labelled hydrogens are shown in red. **b**, NADP $^{+}$  and NADPH labelling patterns (without correction for natural  $^{13}\text{C}$ -abundance) after 48 h incubation with [2,3,3,4,4- $^2\text{H}$ ]glutamine. The indistinguishable labelling of NADP $^{+}$  and NADPH implies lack of NADPH redox active hydrogen labelling. **c**, Pathway diagram showing potential for [2,3,3- $^2\text{H}$ ]aspartate to label NADPH via isocitrate dehydrogenase. **d**, NADP $^{+}$  and NADPH labelling patterns (without correction for natural  $^{13}\text{C}$ -abundance) after 48 h incubation with [2,3,3- $^2\text{H}$ ]aspartate. The indistinguishable labelling of NADP $^{+}$  and NADPH implies lack of redox active hydrogen labelling. **e**, Diagram of [2,3,3,4,4- $^2\text{H}$ ]glutamine metabolism through TCA cycle, tracing labelled

hydrogen. Hydrogen atoms of lighter shade indicate potential H/D exchange with water. **f**, Malate labelling fraction after cells were supplied with [2,3,3,4,4- $^2\text{H}$ ]glutamine for 48 h. **g**, Pathway diagram showing potential for [1,2,3- $^{13}\text{C}$ ]malate (made by feeding [U- $^{13}\text{C}$ ]glutamine) to label pyruvate and lactate via malic enzyme. **h**, Extent of malate and pyruvate/lactate  $^{13}\text{C}$ -labelling. Cells were incubated with [U- $^{13}\text{C}$ ]glutamine for 48 h. M+3 pyruvate indicates malic enzyme flux, which may generate either NADH or NADPH. Similar results were obtained also for M+3 lactate, which was used as a surrogate for pyruvate in cases in which lactate was better detected. The corresponding maximal possible malic enzyme-driven NADPH production rate ranges, depending on the cell line, from  $< 2 \text{ nmol } \mu\text{l}^{-1} \text{ h}^{-1}$  (based on the limit of detection of M+3 pyruvate) to  $6 \text{ nmol } \mu\text{l}^{-1} \text{ h}^{-1}$ . Mean  $\pm$  s.d.,  $n \geq 2$ .



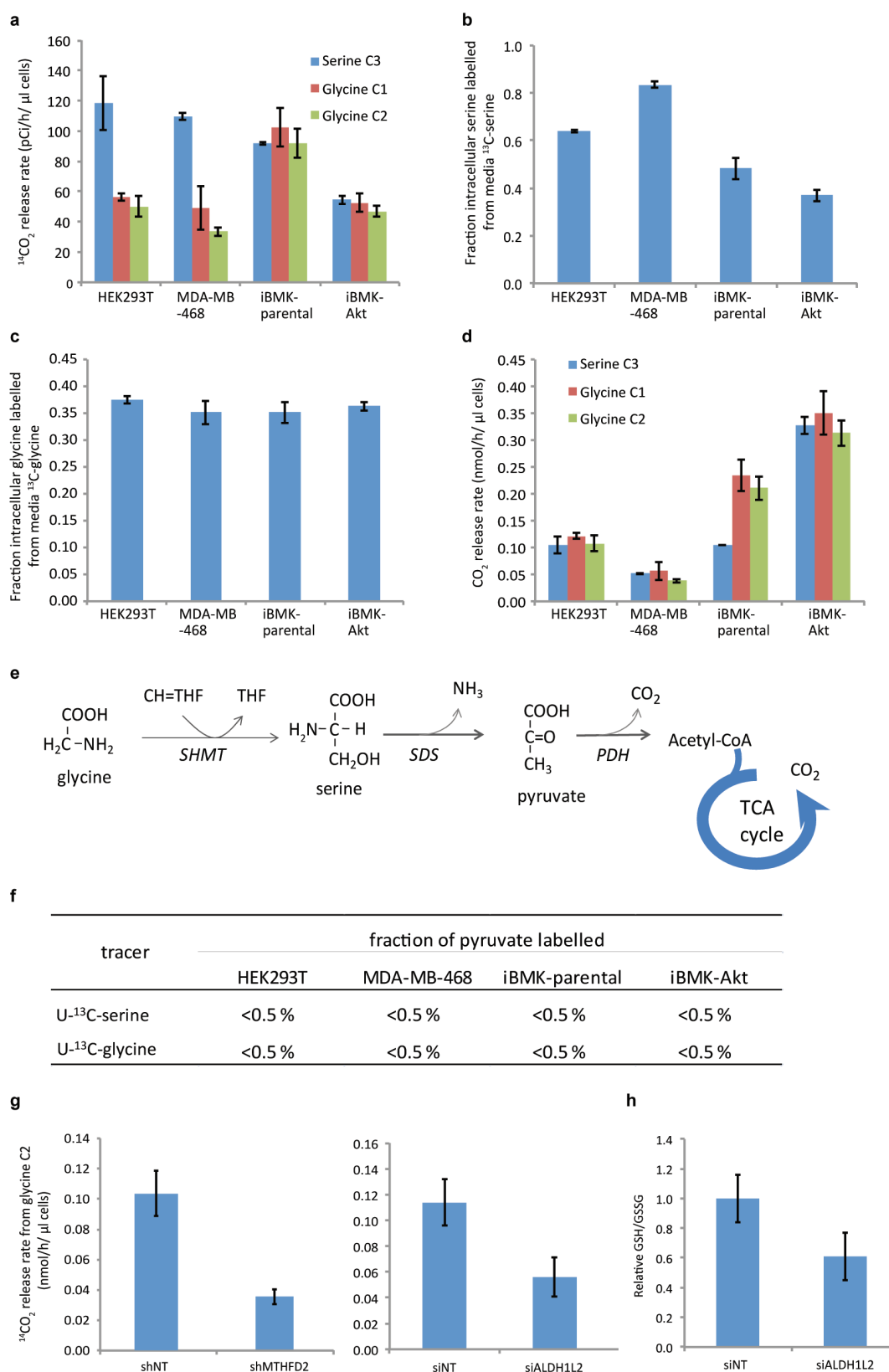
**Extended Data Figure 5 | Computational and experimental evidence for THF-dependent NADPH production.** **a**, Predicted contribution of folate metabolism to NADPH production based on flux balance analysis, using minimization of total flux as the objective function, across different biomass compositions. The biomass fraction of cell dry weight consisting of protein, nucleic acid and lipid was varied as follows: protein 50–90% with a step size of 10%; RNA/DNA 3–20% with step size of 1%, and lipids 3–20% with step size of 1% (considering only those combinations that sum to no more than 100%). With this range of physiologically possible biomass compositions, the model predicts a median contribution of folate metabolism of 24%. Note that with the constraint of experimentally measured biomass composition, yet without constraining the uptake rate of amino acids other than glutamine to be  $\leq 1/3$  of the glutamine uptake rate, the contribution of folate pathway to total NADPH production is predicted to be 23%. **b**, Range of feasible flux through NADPH producing reactions in Recon1 model computed via flux variability analysis under the constraint of maximal growth rate. As shown, the model predicts that each NADPH producing reaction can theoretically have zero flux, with all NADPH production proceeding through alternative pathways. Only reactions

whose flux upper bound is greater than zero are shown. Reactions producing NADPH via a thermodynamically infeasible futile cycle were manually removed. As shown, among all NADPH producing reactions, MTHFD has the highest flux consistent with maximal growth. **c**, Pathway diagram showing potential for [2,3,3-<sup>2</sup>H]serine to label NADPH via methylene tetrahydrofolate dehydrogenase. **d**, NADP<sup>+</sup> and NADPH labelling pattern after 48 h incubation with [2,3,3-<sup>2</sup>H]serine (no glycine present in the media). The greater abundance of more heavily labelled forms of NADPH relative to NADP<sup>+</sup> indicates redox active hydrogen labelling. Results are mean  $\pm$  s.d.,  $n \geq 2$  biological replicates from a single experiment and were confirmed in  $n \geq 2$  experiments. Based on the data in panel **d**, the contribution of MTHFD1 to cytosolic NADPH production spans a broad range (10–40% of total cytosolic NADPH; the range is due to variation across cell lines, experimental noise, and the large KIE<sup>40</sup>). This range includes the flux calculated based on purine biosynthetic rate and <sup>14</sup>CO<sub>2</sub> release from serine (Fig. 3d). Note that the total contribution of the cytosolic folate metabolism to NADPH production can exceed that of MTHFD1, as 10-formyl-THF dehydrogenase also produces NADPH.



**Extended Data Figure 6 | One-carbon units used in purine and thymidine synthesis are derived from serine.** **a**, Serine and ATP labelling pattern after 24 h incubation of HEK293T cells with  $[U-^{13}C]$ serine. The presence of M+1 to M+4 ATP indicates that serine contributes carbon to purines both through glycine and through one-carbon units derived from serine C3. **b**, Quantitative analysis of cytosolic one-carbon unit labelling from measured the intracellular

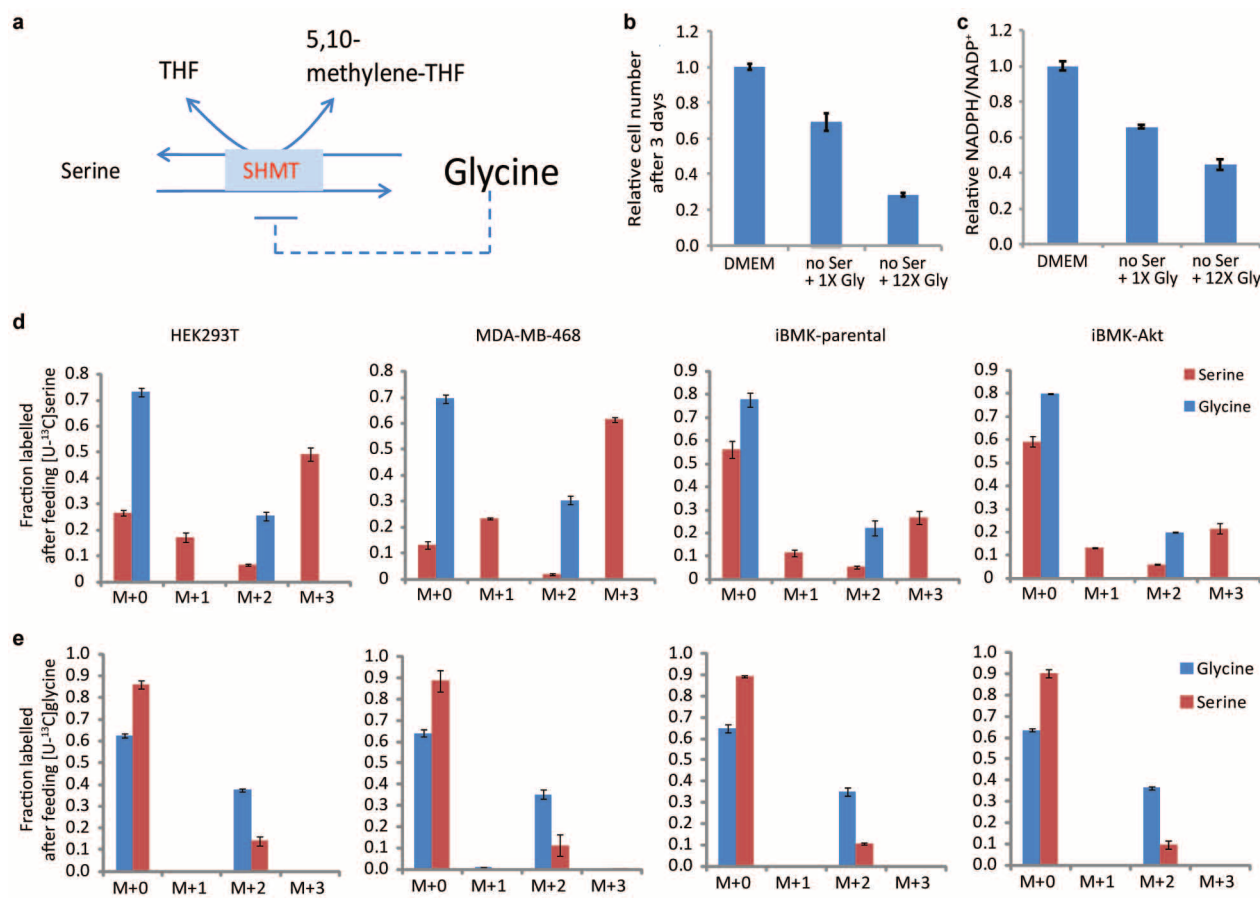
ATP, glycine, and serine labelling reveals that most cytosolic 10-formyl-THF assimilated into purines comes from serine. **c**,  $[U-^{13}C]$ serine labels the methyl group that distinguishes dTTP from UTP. **d**,  $[U-^{13}C]$ glycine does not label dTTP. **e**, The extent of dTTP labelling mirrors the extent of intracellular serine labelling. **f**, Methionine does not label from  $[U-^{13}C]$ glycine. In all experiments, cells were grown in  $[U-^{13}C]$ serine or glycine for 48 h. Mean  $\pm$  s.d.,  $n = 3$ .



**Extended Data Figure 7 | Measurement of CO<sub>2</sub> release rate from serine and glycine by combination of <sup>14</sup>C- and <sup>13</sup>C-labelling.** **a**, <sup>14</sup>CO<sub>2</sub> release rate when cells are supplied with a medium with a trace amount of [3-<sup>14</sup>C]serine, [1-<sup>14</sup>C]glycine or [2-<sup>14</sup>C]glycine. **b**, Fraction of intracellular serine labelled in cells grown in DMEM media containing 0.4 mM [3-<sup>13</sup>C]serine in place of unlabelled serine. The residual unlabelled serine is presumably from *de novo* synthesis. **c**, Fraction of intracellular glycine labelled in cells grown in DMEM medium containing 0.4 mM [U-<sup>13</sup>C]glycine in place of unlabelled glycine.

**d**, CO<sub>2</sub> release rates from serine C3, glycine C1 or C2. **e**, Potential alternative pathway to metabolize glycine or serine into CO<sub>2</sub>, via pyruvate. **f**, Pyruvate labelling fraction after 48 h labelling with [U-<sup>13</sup>C]serine or [U-<sup>13</sup>C]glycine. The lack of labelling in pyruvate indicates that serine and glycine are not metabolized through this pathway. **g**, Knockdown of *MTHFD2* or *ALDH1L2* decreases CO<sub>2</sub> release from glycine C2. **h**, Knockdown of *ALDH1L2* decreases the GSH/GSSG ratio. Mean ± s.d., *n* = 3.

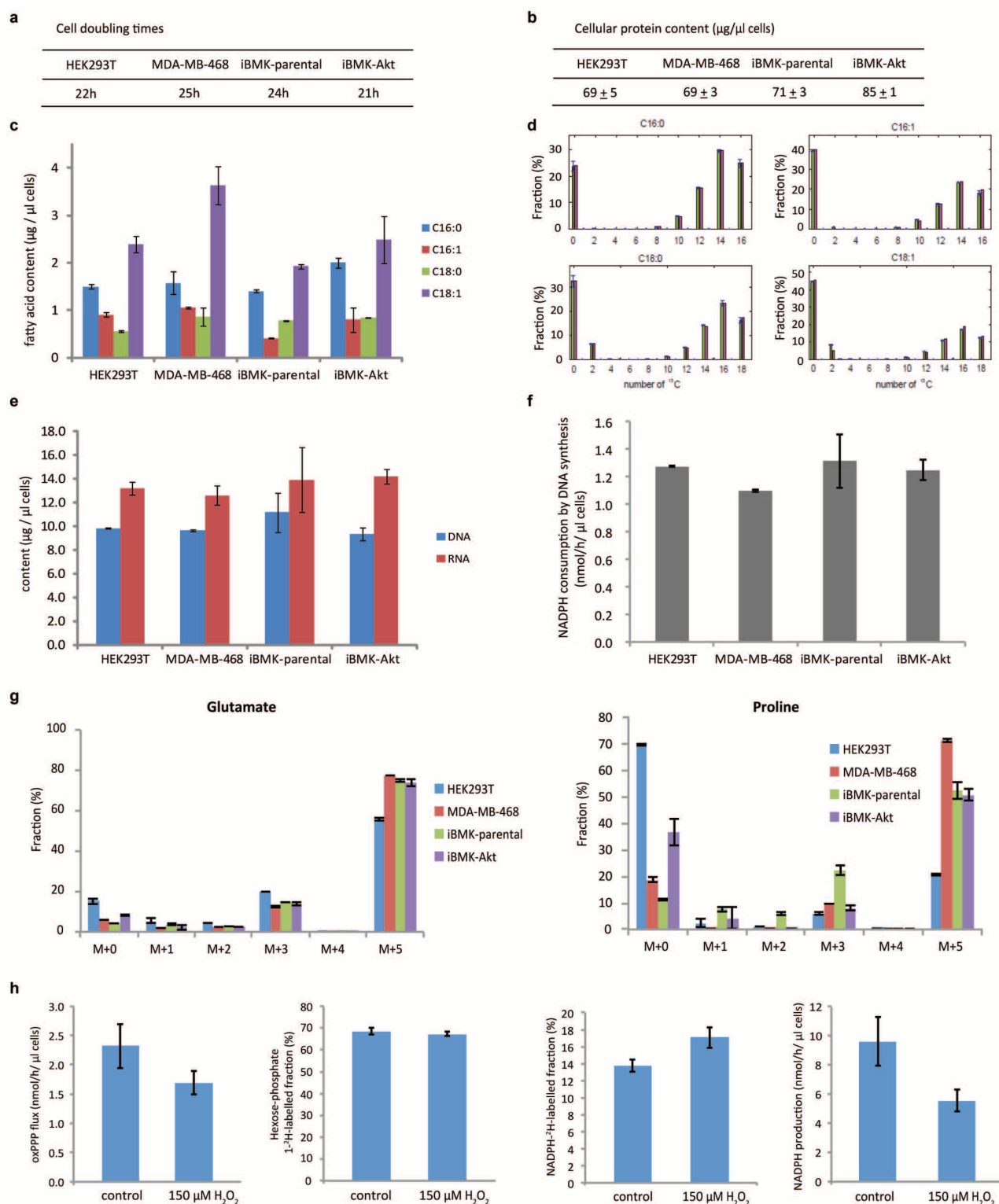




**Extended Data Figure 8 | In the absence of serine, elevated concentrations of glycine inhibit cell growth and decrease the NADPH/NADP<sup>+</sup> ratio.**

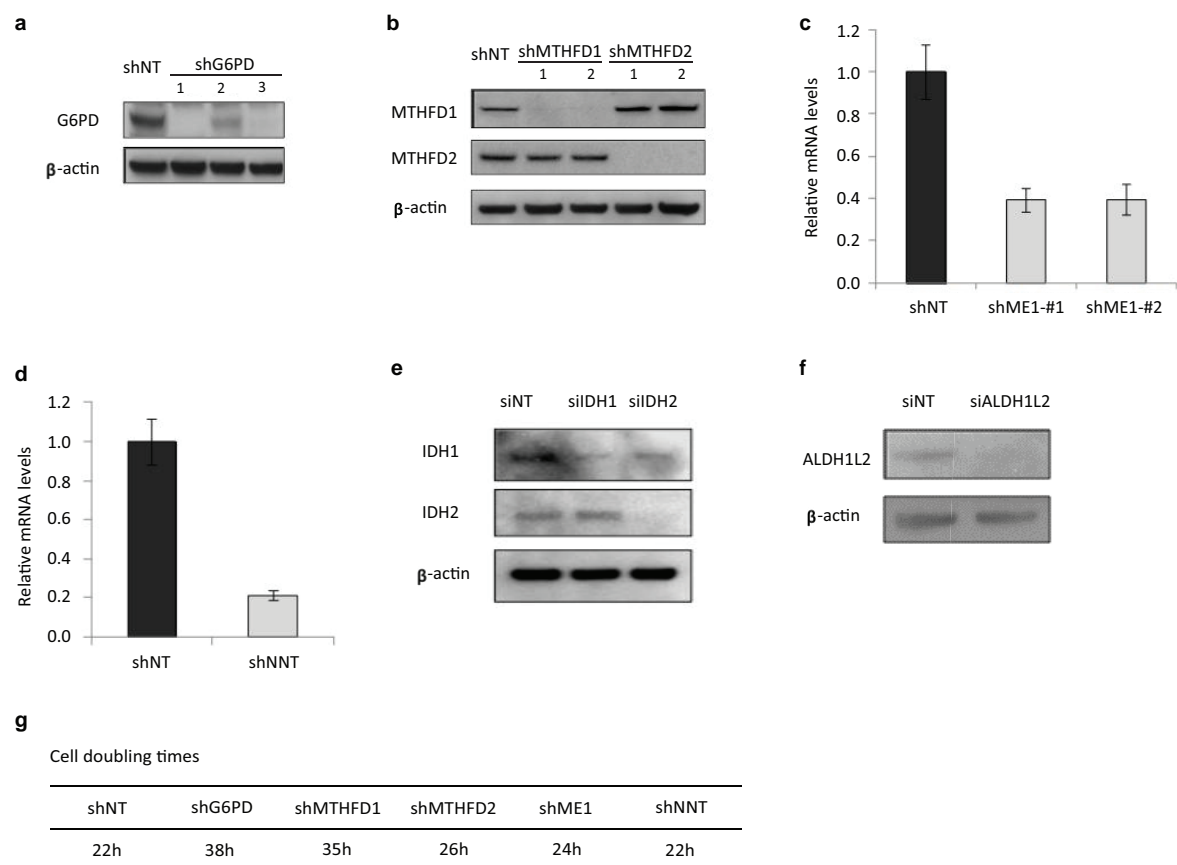
**a**, Schematic of serine hydroxymethyltransferase reaction. High glycine may either inhibit forward flux (product inhibition) or drive reverse flux. **b**, Relative cell number after culturing HEK293T cells for 3 days in regular DMEM, DMEM with no serine or DMEM with no serine and 12.5-times the normal

concentration of glycine (5 mM instead of 0.4 mM). **c**, Relative NADPH/NADP<sup>+</sup> ratio (normalized to cells grown in DMEM) after culturing HEK293T cell for 3 days in regular DMEM, DMEM with no serine or DMEM with no serine and 12.5-times the normal concentration of glycine. **d**, **e**, Labelling of serine and glycine after feeding [U-<sup>13</sup>C]serine or [U-<sup>13</sup>C]glycine reveals reverse serine hydroxymethyltransferase flux. Mean  $\pm$  s.d.,  $n = 3$ .



**Extended Data Figure 9 | Quantitative analysis of NADPH consumption for biomass production and antioxidant defence.** **a**, Cell doubling times, which are inversely proportional to biomass production rates. **b**, Cellular protein content. **c**, Cellular fatty acid content (from saponification of total cellular lipid). **d**, Quantification of fatty acid synthesis versus import, with synthesis but not import requiring NADPH. HEK293T cells were cultured in  $[\text{U}-^{13}\text{C}]$ glucose and  $[\text{U}-^{13}\text{C}]$ glutamine until pseudo-steady state, and fatty acids saponified from total cellular lipids and their labelling patterns measured (green bars), and production versus import of each fatty acid was stimulated based on this experimental data. The fractional contribution of each route was determined by least square fitting, with the theoretical labelling pattern based on the elucidated routes shown (pink bars). Similar data were obtained also for

MD-MBA-468, iBMK-parental, and iBMK-Akt cells (not shown) and used to calculate associated NADPH consumption by fatty acid synthesis. **e**, Cellular DNA and RNA contents. **f**, NADPH consumption by *de novo* DNA synthesis. **g**, Proline and glutamate labelling patterns after 24 h in  $[\text{U}-^{13}\text{C}]$ glutamine media, which was used to quantitate different proline synthesis routes and associated NADPH consumption. **h**, Quantitative analysis of cytosolic NADPH consumption in normally growing HEK293T cells (control) and non-growing cell under oxidative stress ( $150 \mu\text{M}$   $\text{H}_2\text{O}_2$ , 5 h). Total cytosolic NADPH turnover was measured based on the absolute oxidative pentose phosphate pathway flux divided by the fractional contribution of the oxidative pentose phosphate pathway to total NADPH as measured using  $[\text{2H}]$ NADPH formation from  $[\text{1-2H}]$ glucose. Mean  $\pm$  s.d.,  $n = 3$ .



**Extended Data Figure 10 | Confirmation of knockdown efficiency by western blot or qPCR. a,** Western blot for G6PD knockdown. **b,** Western blot for MTHFD1 and MTHFD2 knockdown. **c,** mRNA level for *ME1* knockdown. **d,** mRNA level for *NNT* knockdown. **e,** Western blot for IDH1 and IDH2 knockdown. **f,** Western blot for ALDH1L2 knockdown. **g,** Cell doubling times of HEK293T with stable knockdown of indicated genes (results for different hairpins of the same gene were indistinguishable).

# CAREERS

@NATUREJOBS Follow us on Twitter for the latest news and features [go.nature.com/e492gf](http://go.nature.com/e492gf)

NATUREJOBS BLOG The latest on science-careers news and tips [go.nature.com/ielkkf](http://go.nature.com/ielkkf)

NATUREJOBS For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)



GARY WATERS/IKON IMAGES/GETTY

## COLUMN

### A test that fails

A standard test for admission to graduate school misses potential winners, say **Casey Miller** and **Keivan Stassun**.

Universities in the United States rely too heavily on the graduate record examinations (GRE) — a standardized test introduced in 1949 that is an admissions requirement for most US graduate schools. This practice is poor at selecting the most capable students and severely restricts the flow of women and minorities into the sciences.

We are not the only ones to reach this conclusion. William Sedlacek, professor emeritus of education at the University of Maryland, College Park, who has written extensively on the issue, notes that studies find only a weak correlation between the test and ultimate success in science, technology, engineering and maths (STEM) fields. De-emphasizing the GRE and augmenting admissions procedures with measures of other attributes — such as drive, diligence and the willingness to take scientific risks — would not only make graduate admissions more predictive of the ability to do well but would also increase diversity in STEM.

#### TEST DISPARITIES

The GRE, like most standardized tests, reflects certain demographic characteristics of test-takers — such as family socioeconomic status — that are unrelated to their intellectual capacity or academic preparation. The exam's 'quantitative score' — the portion measuring maths acumen, which is most commonly scrutinized in admissions to STEM PhD programmes — correlates closely with gender and ethnicity (see 'The great divide'). The effect is powerful. According to data from Educational Testing Service (ETS), based in Princeton, New Jersey, the company that administers the GRE, women score 80 points lower on average in the physical sciences than do men, and African Americans score 200 points below white people. In simple terms, the GRE is a better indicator of sex and skin colour than of ability and ultimate success.

These correlations and their magnitude are not well known to graduate-admissions committees, which have a changing rota of faculty members. Compounding the problem, some admissions committees use minimum GRE scores to rapidly filter applications; for example, any candidate scoring below 700 on the 800-point quantitative test section may be discarded. Using GRE scores to filter applicants in this way is a violation of ETS's own guidelines.

This problem is rampant. If the correlation between GRE scores and gender and ethnicity is not accounted for, imposing such cut-offs adversely affects women and minority applicants. For example, in the physical sciences, ►



► only 26% of women, compared with 73% of men, score above 700 on the GRE Quantitative measure. For minorities, this falls to 5.2%, compared with 82% for white and Asian people.

The misuse of GRE scores to select applicants may be a strong driver of the continuing under-representation of women and minorities in graduate school. Indeed, women earn barely 20% of US physical-sciences PhDs, and under-represented minorities — who account for 33% of the US university-age population — earn just 6%. These percentages are striking in their similarity to the percentage of students who score above 700 on the GRE quantitative measure.

Why is the GRE misused? Admissions committees are busy, and numerical rankings are easy to sort. We believe that faculty members also often presume that higher scores imply that the test-taker has a greater ability to become a PhD-level scientist. Yet research by ETS indicates that the predictive validity of the GRE tests is limited to first-year graduate-course grades, and even that correlation is meagre in maths-intensive STEM fields.

Why should graduate-admissions committees care about fixing the problem? First, diversity, in the form of individuals with different perspectives, backgrounds and experiences, is a key component of innovation and problem solving, a concept that business and industry have come to recognize. Less diversity in STEM graduate programmes means slower progress in tackling today's scientific and technical challenges. Second, the overall PhD completion rate in US STEM graduate programmes is a disappointing 50%. Although graduate programmes certainly produce successful students who continue on to productive science careers, we think that many faculty members would agree that such a low PhD completion rate is a poor return on the investment in recruiting and training students. Indeed, STEM graduate programmes are failing not only from the diversity standpoint, but also from a success standpoint.

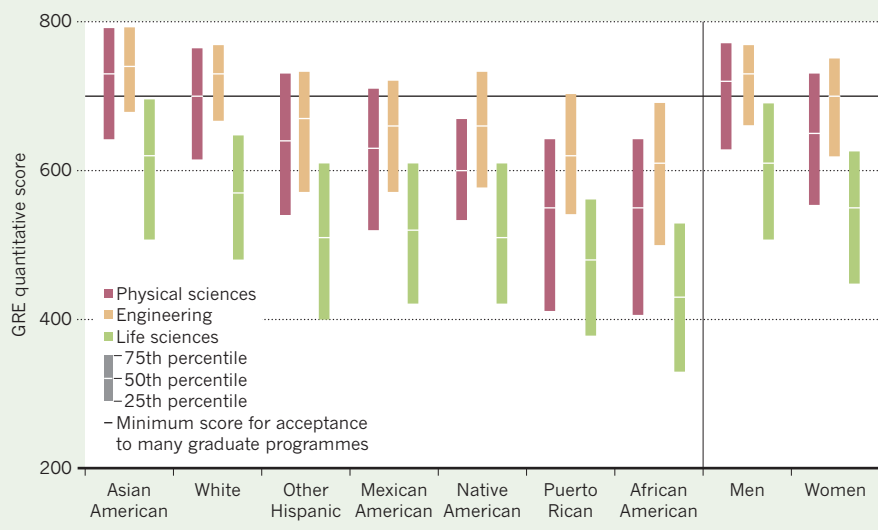
## ALTERNATIVE SELECTION

So what should universities do? Instead of filtering by GRE scores, graduate programmes can select applicants on the basis of skills and character attributes that are more predictive of doing well in scientific research and of ultimate employability in the STEM workforce. Appraisers should look not only at indicators of previous achievements, but also at evidence of ability to overcome the tribulations of becoming a PhD-level scientist.

A few innovative PhD programmes, including the bridge programmes at the University of South Florida in Tampa and Fisk–Vanderbilt in Nashville, Tennessee (in which we are involved) are doing this. They have achieved completion rates above 80%, well above the national average, and are greatly boosting participation by women and minorities (see *Nature* **504**, 471–473; 2013). The admissions process includes an

## THE GREAT DIVIDE

The data represent the scores typically achieved in the quantitative reasoning test of the graduate record examinations (GRE) by US students from different ethnic groups applying for graduate school. In the physical sciences, a minimum score of 700 is required by many PhD programmes.



SOURCE: ETS

interview that examines college and research experiences, key relationships, leadership experience, service to community and life goals. The result is a good indication of the individual's commitment to scientific research and a good assessment of traits such as maturity, perseverance, adaptability and conscientiousness atop a solid academic foundation. The combination of academic aptitude and these other competencies points to the likelihood of high achievement in graduate school and in a STEM career.

How have the students admitted to these courses performed? In the Fisk–Vanderbilt programme, 81% of the 67 students who have entered the programme — including 56 under-represented minorities and 35 women — have earned, or are making good progress towards, their PhDs. And all students who have completed PhDs are employed in the STEM workforce as postdocs, university faculty members or staff scientists in national labs or industry. From the standpoint of optimal outcomes — earning a PhD and obtaining employment in the STEM workforce — the GRE has proved irrelevant. Indeed, 85% of these young scientists would have been eliminated from consideration for PhD programmes by a GRE quantitative cut-off score of 700.

The only downside is that interviews take about 30 minutes each. But the number of interviews need not be large, and the tremendous insight garnered justifies the time. ETS is even marketing a tool for referees to evaluate applicants' personal attributes. The company developed it in part as a response to calls from applicants and graduate programmes for alternative measures of student potential for long-term achievement that is not captured by GRE.

We often hear admissions committee members say, 'We would admit women and minorities if they were qualified'. This mindset reflects

long-standing admissions practices that systematically, if inadvertently, filter out women and minorities. At the same time, these practices are no better than a coin flip at identifying candidates with the potential — and the mettle — to earn a PhD.

Let us be frank: we believe that many STEM faculty members on admissions committees and upper-level administrators hold a deep-seated and unfounded belief that these test scores are good measures of ability, of potential for doing well in graduate school and of long-term potential as a scientist, and that students who score poorly on standardized exams are not likely to become PhD-level scientists. These assumptions are false.

This is not a call to admit unqualified students in the name of social good. This is a call to acknowledge that the typical weight given to GRE scores in admissions is disproportionate. If we diminish reliance on GRE and instead augment current admissions practices with proven markers of achievement, such as grit and diligence, we will make our PhD programmes more inclusive and will more efficiently identify applicants with potential for long-term success as researchers. Isn't that what graduate school is about? ■

**"In simple terms, the GRE is a better indicator of sex and skin colour than of ability and ultimate success."**

**Casey Miller** is an associate professor in the physics department at the University of South Florida in Tampa. **Keivan Stassun** is a professor in physics and astronomy at Vanderbilt University and Fisk University in Nashville, Tennessee.

# FACE IN THE DARK

*Those who came before.*

BY DEBORAH WALKER

Conglomerate time ticked. In the 24-hour regime of the habitat, the automated lights faded to simulated night. The night blinds shuttered into place. Throughout the Conglomerate colonies, time is unified and indifferent to the rotation of local suns.

Dan lay in the darkness counting the deep, slow breaths of his wife, drifting towards the pull of sleep, until ... he saw it. He shuddered.

"What's the matter?" Arrelle whispered.

"I thought you were asleep."

"No."

The bed was small, but she seemed too far away. Dan slipped closer and felt the warmth of her body.

"Do you see one?" she asked.

"It's by the window." Dan pointed to the shadows, at the figure resolved out of darkness. The indistinct, small shape walked with jerky movements like a worn recording made on an antiquated device.

"I see it," said Arrelle, her voice devoid of inflection. "The doctor ran the tests again, today. He says that the air's clean. There's no trace of hallucinogens."

The shadow figure moved a few centimetres.

"I'm beginning to think that we've all gone mad," said Arrelle.

Dan said nothing.

"Did you hear me, Dan?"

Dan hesitated before saying: "Father McConnell will try another exorcism tomorrow."

Arrelle sighed. "It won't do any good."

McConnell, the colony's trans-faith priest, had read the *Conglomerate Book of Devotion* blank trying different exorcisms. All human things have value. Faith was levelled into a matter of equality. A Catholic exorcism was equal to a Voodoo prayer chant or to a Vedic mantra. This had proved to be the case. Each of McConnell's expulsions had had the same result: failure.

"What do you think it wants, Dan?"

"I ... don't know." Can the dead want anything? "Perhaps it's just a *restligeist*." That

was the most popular theory in the mining labs. "It's just an echo. Something that's ingrained into the air

or the stones of this planet."

"The Stone Tape theory? You think they're alien memories somehow imprinted onto the wall. A psychic echo?" asked Arrelle. She looked at the shadow creeping over the wall. "How can that be true? It doesn't make any sense. The Conglomerate doesn't recognize psychic phenomena."



"I know," said Dan. "But they've run the tests. There's no electromagnetic disturbances, or changes in ionization, or changes in the radiation levels. There's nothing we can detect."

Arrelle touched Dan's arm. "Why can't you do something? You work in the labs. Why don't you do something?"

"I'm a *geologist*, Arrelle," said Dan. "The medical and the physics teams are doing everything they can." In fact, it seemed that most of the conversations in the mining labs were devoted to speculations about the night visitors. The labs were lagging behind the Conglomerate's schedules. Dan was a supervisor, but he had no idea how he could motivate his team. The colony was slowly falling apart because of these night shadows.

Arrelle pulled the sheets around her neck. "I'm so cold."

"It's all in your imagination." Dan tapped a command into his wrist-bracelet. He held the small screen towards her, lighting her face with flickering chiaroscuro fluorescence. "Look," he said. "Look. There's no change in temperature."

"Then why am I so cold?"

The shadow figure slowly turned. It raised its three arms above its head and slowly sketched out an incomprehensible gesture.

"I read the Conglomerate's first contact protocols again," said Arrelle.

"How can we make first contact with that?"

"I want to try, Dan. We can't live like this."

"And what would you say?"

"I'd say that we were sorry."

A small oval of red light formed in the shadow's face. "Its eye is open," said Dan. "That means that it'll be gone soon." Its red eye hung in the sky of its face like an accusation.

Dan and Arrelle watched as the shadow moved along the wall. When it reached the corner it faded out of sight.

"I wish we could leave," said Arrelle. She curled lower into the sheets, folding her body, knees tight to her chest.

"You know that we can't," said Dan. It was impossible to leave a Conglomerate colony. Once the seed ship had deposited its frozen bodies, it moved on, following the scout ships along the path of Conglomerate expansion. Conglomerate colonization was an efficient, automated process. The colonists slowly melted

into re-life, awakening to a colony made safe and habitable by the soldier and construction auto-drones.

Dan sighed. He got out of bed, reaching for his robe. Whatever the computer said, it was cold in this room. He walked over to the window and touched a button to iris open the blinds. Next time he saw the thing, he'd keep quiet. Talking about it with Arrelle was only making things worse. Whatever the shadow was, there was nothing Dan could do about it.

The low red-eye sun, slung in the sky, cast long shadows. In the distance, against the rose-red light, the towers of the ruined city were bone white needles, extending towards the horizon. It must have once been marvellous. Dan saw a discarded Conglomerate Humvee buggy. He saw the remains of a soldier auto-drone. He saw the fields of three armed skeletons, bones large and small, slowly dissolving in the acid atmosphere.

"It's not our fault," whispered Dan. "We didn't know."

"We never asked," said Arrelle. "We just accepted the Conglomerate's offer."

Across the horizon the shadows moved, remnants of the evicted dead who were loathe to leave their home. ■

*Find Deborah in the British Museum trawling the past for future inspiration.*